

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: March 12, 2005, 13:08:19 ; Search time 887 Seconds
(without alignments)
7055.073 Million cell updates/sec

Title: US-10-627-132-30
Perfect score: 5648
Sequence: 1 MEASGLVAGSHRNELVLI.....IDPFIVRTKGPDPVRCGNC 1052

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5537552 seqs, 2974263231 residues
Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10627132/runat_04032005_113220_21246/app_query.fasta_1.1223
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNIT8=bits -START=1 -END=-1 -MATRIX=BLOSUM62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MOD=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10627132 @CGN 1 1 463 @runat_04032005_113220_21246
-NCPU=6 -ICPU=3 -NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DISPLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PTC_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PTCUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	
1	5648	100.0	3443	17	Sequence 29, Appl Sequence 29, Appl	
2	5520	97.7	3589	17	Sequence 24930, A Sequence 24930, A	
3	5347.5	94.7	3448	18	Sequence 15093, A Sequence 15093, A	
c	4	5055	3028	16	Sequence 29, Appl Sequence 29, Appl	
5	4788.5	84.8	3517	9	Sequence 13, Appl Sequence 13, Appl	
6	4561	80.8	2830	18	Sequence 139789, Sequence 139789, Sequence 139789	
7	4477	79.3	3851	17	Sequence 69, Appl Sequence 69, Appl	
8	4395	77.8	2687	17	Sequence 833, Appl Sequence 833, Appl	
9	4091	72.4	3776	9	Sequence 3, Appl Sequence 3, Appl	
10	4089	72.4	3505	17	Sequence 31130, A Sequence 31130, A	
11	4089	72.4	3783	18	Sequence 86998, A Sequence 86998, A	
12	4073.5	72.1	3936	9	Sequence 9, Appl Sequence 9, Appl	
13	4073.5	72.1	3969	16	Sequence 49, Appl Sequence 49, Appl	
14	4073.5	72.1	3969	16	Sequence 17, Appl Sequence 17, Appl	
15	4073.5	72.1	3969	16	Sequence 37, Appl Sequence 37, Appl	
16	4073.5	72.1	3969	17	Sequence 49, Appl Sequence 49, Appl	
17	4073	72.1	4297	18	Sequence 44334, A Sequence 44334, A	
c	18	4059	71.9	4225	18	Sequence 72205, A Sequence 72205, A
19	4005	70.9	3568	16	Sequence 41, Appl Sequence 41, Appl	
20	4005	70.9	3568	16	Sequence 1, Appl Sequence 1, Appl	
21	4005	70.9	3568	17	Sequence 41, Appl Sequence 41, Appl	
22	3982.5	70.5	3814	18	Sequence 45185, A Sequence 45185, A	
23	3974.5	70.4	3786	9	Sequence 7, Appl Sequence 7, Appl	
24	3974.5	70.4	3813	16	Sequence 17, Appl Sequence 17, Appl	
25	3974.5	70.4	3813	16	Sequence 25, Appl Sequence 25, Appl	
26	3974.5	70.4	3813	16	Sequence 45, Appl Sequence 45, Appl	
27	3974.5	70.4	3813	17	Sequence 17, Appl Sequence 17, Appl	
28	3974.5	70.4	3987	18	Sequence 45192, A Sequence 45192, A	
29	3919.5	69.4	3962	18	Sequence 72402, A Sequence 72402, A	
30	3917.5	69.4	3812	18	Sequence 101245, Sequence 101245, Sequence 101245	
31	3914.5	69.3	3799	16	Sequence 5, Appl Sequence 5, Appl	
32	3914.5	69.3	3799	16	Sequence 21, Appl Sequence 21, Appl	
33	3914.5	69.3	3799	17	Sequence 21, Appl Sequence 21, Appl	
34	3912.5	69.3	3746	16	Sequence 29, Appl Sequence 29, Appl	
35	3912.5	69.3	3746	16	Sequence 49, Appl Sequence 49, Appl	
36	3912.5	69.3	3773	16	Sequence 9, Appl Sequence 9, Appl	
37	3912.5	69.3	3773	16	Sequence 5, Appl Sequence 5, Appl	
38	3912.5	69.3	3773	17	Sequence 9, Appl Sequence 9, Appl	
39	3903.5	69.1	3704	16	Sequence 13, Appl Sequence 13, Appl	
40	3903.5	69.1	3704	16	Sequence 57, Appl Sequence 57, Appl	
41	3903.5	69.1	3704	17	Sequence 13, Appl Sequence 13, Appl	
42	3903.5	69.1	3744	18	Sequence 100413, Sequence 100413, Sequence 100413	
43	3901	69.1	3614	15	Sequence 9, Appl Sequence 9, Appl	
44	3899	69.0	3725	16	Sequence 45, Appl Sequence 45, Appl	
45	3899	69.0	3725	16	Sequence 13, Appl Sequence 13, Appl	

ALIGNMENTS

RESULT 1
US-10-627-132-29
; Sequence 29, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhuga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE OF INVENTION: Thereof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31

```
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 3443
; TYPE: DNA
; ORGANISM: Zea mays
US-10-627-132-29

Alignment Scores:
Pred. No.: 0 Length: 3443
Score: 5648.00 Matches: 1052
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-10-627-132-30 (1-1052) x US-10-627-132-29 (1-3443)

Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
Db 92 ATGAGGCGCAGCGCGGCTGCTGCGCGCTCGCACACCGAAGAGCTGGTCTGATC 151
Qy 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGluIleCys 40
Db 152 CGGGGCGCACAGGACCCCAAGCCGCTGCGGCGCTGAGCGGCGAGGTGTCGAGATATGC 211
Qy 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
Db 212 GGCAGCAGAGGTCGGGCTCACCGTGGACGGCGACCTCTTCGTCGCTGCAAGAGTGGCGG 271
Qy 61 PheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsnCysPro 80
Db 272 TTCCTCCGCTGTCGGGCTCTACGAGTACGAGCGCGGAGGAGGACGACGAACTGCCCC 331
Qy 81 GlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAspAsp 100
Db 332 CAGTGCMAAGACGCGCTACAAGCGCTCAAGGGAGCGCGAGGTTGCGGGGACGATGAC 391
Qy 101 GluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnArg 120
Db 392 GAGGAGGACATCGACGACTCGGAGCACGAGTTCAACATCGACGACGAGAAATCAGCAGAG 451
Qy 121 GlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMet 140
Db 452 CAGCTGGAGGCGCAACATCGACAGACAGCCAGATCACCGAGGCGATGCTGCACGCGAGATG 511
Qy 141 SerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIleProProlle 160
Db 512 AGCTACGGGAGGGGCGCGCGACGCGCGACGCGCAACACACCGCGCAGATCCCGCCCATC 571
Qy 161 IleThrGlySerArgSerValProValSerGlyGluPheProlleThrAsnGlyTyrGly 180
Db 572 ATCACCGGCTCCCGCTCCGCTCGGCTGAGCGGTGAGTTTCGATATACCACGGGTATGGC 631
Qy 181 HisGlyGluValSerSerLeuHisLysArgIleHisProTyrProValSerGluPro 200
Db 632 CACGGCGAGGTCTCTCTCTCCCTGCACAAAGCGCATCCATCCGTAACCTGTGTCTGAGCCA 691
Qy 201 GlySerAlaLysTrpAspGluLysLysGluValSerTrpLysGluArgMetAspAspTrp 220
Db 692 GGGAGTGCACAGTGGGCGAGAGAAAGAGTGTGAGCTGGAAGGAGGAGGATGACGACTGG 751
Qy 221 LysSerLysGlnGlyIleGlyGlyValAlaAspProGluAspMetAspAlaAspVal 240
Db 752 AAGTCCAGCAGGGCATCTCCGGCGCGCGCGCGATCCCGAAGACATGGACGCCACGCTG 811
Qy 241 AlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSerSerLys 260
Db 812 GCATCTAAGCAGCAGCGCGAGGCGAGCGCTGTGAGGAAGGTGTCCATCGCGTTCGAGCAAG 871
Qy 261 ValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPhePheLeuArg 280
Db 872 GTGAACCGGTACCGGATGGTGTGTCGTGTCGTCTCGTTGTGTCGTCTCTCTCTCTCTCCGG 931
```

```
Qy 281 TyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerIleIleCys 300
Db 932 TACCGTATCTGACACCCCGTCCCGAGCGCATCGGGCTGTGGCTCGTCTCCATCATCTGC 991
Qy 301 GluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPheProlleAsp 320
Db 992 GAGATCTGGTTCGCCATCTCTGGATCTCTGACACCGTTCGCCAAGTGGTTCGCCATCGAC 1051
Qy 321 ArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGluProSerLeu 340
Db 1052 CGCGAGCAGTACCTCGACCGCTCTCCCTCAGGTACGAGAGGAGGAGCGCTGCTCTG 1111
Qy 341 LeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluProProLeuVal 360
Db 1112 CTGTCCGGCGGTGGACCTGTTCGTGAGCACCGTGGACCGCTCAAGGAGCGCGCTGGTG 1171
Qy 361 ThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCys 380
Db 1172 ACCGCCAACACACCGTGTCTCCATCTCGCGTAGACTACCCCGTGGACAAAGGTCTCTCG 1231
Qy 381 TyrValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGlu 400
Db 1232 TACGTCTCTCCAGCAGCGCGGTGATGCTGACGTTCAGTCTGCTGTCCGAGACGCGCGAG 1291
Qy 401 PheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArgAlaProGlu 420
Db 1292 TTCGGCGCGCAAGTGGTGGCTTCTGCAAGAAGTTCGGCATCGAGCCCGCGCGCGAG 1351
Qy 421 PheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThrPheValGln 440
Db 1352 TTCTACTTCTCGCTCAAGGTGCGACTACTCAAGGACAAAGGTGAGCCCGCTTCGTGTCAG 1411
Qy 441 GluArgArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAlaLeuVal 460
Db 1412 GAGCGCGCGCCCATGAAGAGAGAGATGATGAGAGTTCAAGGTCCGATCAACGCGCTGGTG 1471
Qy 461 AlalysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrProTrpPro 480
Db 1472 GCCAAGGCCATGAAGGTGCGCGCAGAGGGTGGATCATGAAGGACGCGCACCGCTGGCCC 1531
Qy 481 GlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGly 500
Db 1532 GGGAAACAACCCCGCAGCACCCCGCGCATGATCAGGTGTTCCTGGGCGCACAGCGGGCGG 1591
Qy 501 HisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlyLysArgPro 520
Db 1592 CACGACACCGAGGGCAACAGAGCTGCCCGCTCTGTGTACGTCTCCCGTGAAGACGCGCG 1651
Qy 521 GlyPheGlnHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSerAlaVal 540
Db 1652 GGATTCAGCACCCACAAAGAGCGCGCCCATGAACGCTCTGATTCGCGTCTCCGCGGTG 1711
Qy 541 LeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLys 560
Db 1712 CTGACCAACCGCGCATTCATGCTCAACTGGACTGTGATCACTACATCAACACAGCAAG 1771
Qy 561 AlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysValCysTyr 580
Db 1772 GCCATCCGGGAGGCGATGTGCTTCTCATGACCCCTCAGGTCCGCGGAGAGGTCTGCTAC 1831
Qy 581 ValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsnArgAsn 600
Db 1832 GTTCAGTTCGCCGAGAGGTTCGACGGCATCGACGTGCACGACCGCATCGCTAACACAGGAAC 1891
Qy 601 ThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProValTyrVal 620
Db 1892 ACCGTCTTCTTCGACATCAACATGAAGGGGTGACCGCATCCAAGGGCCCCGTGTACGTC 1951
Qy 621 GlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProProLysGlyPro 640
Db 1952 GGGAGAGGTCGTGTTCCGGCGCGCAGCGCTCTACGGCTACACACCTTCCCAAGGGACCC 2011
```


QY 641 LysArgProLysMetValThrCysAspCysCysProCysPheGlyArgLysIleArgLys 660
DB 2012 AAGAGCCCAAGATGATGACCTGCGACCTGCTCCCGCTTTCGCGCGAAGAGCGAAA 2071
QY 661 HisAlaLysAspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAspLysGlu 680
DB 2072 CAGGCCAAGGACGGGCTCGCGAGGGCACCGCTGATATGGAGTAGATAGCAAGGAG 2131
QY 681 MetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThr 700
DB 2132 ATGCTCATGCTCCACATGAATTCAGAAAGCGGTTTCGGGCGATCGCGCGGTCGTCACG 2191
QY 701 SerThrLeuMetGluGluGlyValProSerSerSerSerProAlaAlaLeuLys 720
DB 2192 TCGACGCTGATGAGAGAGCGCGGCTCTCTTCGTGAGCGCCCGCGCTCTCTCAAG 2251
QY 721 GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeu 740
DB 2252 GAGGCCATCCATGTCATCAGCTGCGGTACGAGGACAGACCGACTGGGGCTGGAGCTG 2311
QY 741 GlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg 760
DB 2312 GGGTGGATCTACGGTTCGATCAGGAGGACATCTCTGACGGGTTCAAGATGCACTGCCGC 2371
QY 761 GlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIle 780
DB 2372 GGGTGGGCTCGGTACTGATGCGAGCGGCGGCTTCAAGGGGTGCGCGCGATC 2431
QY 781 AsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePhe 800
DB 2432 AATCTATCGGACCGTCTCAACAGAGTGTCCGGTGGCGCTGGGTCTCGTCAAGATCTTC 2491
QY 801 PheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGlu 820
DB 2492 TTCAGCGCGCACAGCGCCCTGCTGTACGCGTACAAAGACGCGCAACTCAAGTGGCTGGAG 2551
QY 821 ArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyr 840
DB 2552 CGCTTCGGCTTACATCAACACCACTTACCCCTTACCTCGCTCCCGCTGCTCGCCCTAC 2611
QY 841 CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr 860
DB 2612 TGCACCTCCCGCGCTGCTGCTCTCACCGCAAGTTTCATCTGCCGTGATAGCAG 2671
QY 861 PheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGlu 880
DB 2672 TTCGCCAGCCTCTTCTCATCGCCCTTTCATGTCTCATCTTCGCGACGGCATCTCGGAG 2731
QY 881 MetArgTrpSerGlyValSerIleGluTrpTrpArgAsnGluGlnPheTrpValIle 900
DB 2732 ATGCGGTGGAGGGGTGAGCATCGAGAGTGTGGAGAAACGACAGCTTCTGGGTCTATC 2791
QY 901 GlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLysValLeuAlaGly 920
DB 2792 GCGGCGGTGTCGCGCATCTTTCGCGCTGTCGAGGGCTGCTCAAGGTCTCTCGCCGG 2851
QY 921 IleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPheAla 940
DB 2852 ATCGACCACTTCAACGCTCACTTCAAGGCCACCGCGGACGAGGACGAGTTCGCTGCC 2911
QY 941 GluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeuLeuIle 960
DB 2912 GAGCTCTACGCTTCAAGTGGACCACTGCTCTCTCATCCGCGCCACACGCTGCTCATCAT 2971
QY 961 AsnValIleGlyValValAlaGlySerPheAlaIleAsnAsnGlyTyrGlnSerTrp 980
DB 2972 AACGTATCGGCTGTCGCGCGCATCTCCGAGCGCATCAACAGGGTACCGAGTCTCTGG 3031
QY 981 GlyProLeuPheGlyLysLeuPheAlaPheTrpValIleValHisLeuTyrProPhe 1000
DB 3032 GGGCCCTCTTCGCGAAGCTCTTCTCGCTTCTCGGTTCATGCTCCACTCTACCCGTTTC 3091
QY 1001 LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIleTrpSerIle 1020

DB 3092 CTCAAGGGGCTCATGGGCGCGAGAACAGGACGCCACCGTTGTGTCTATCTGGTCCATT 3151
QY 1021 LeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThr 1040
DB 3152 CTGCTGGGCTCCATCTTCTCTCTGCTGCTGGGTGAGATCGACCTTTTCATCTGTCAAGACC 3211
QY 1041 LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
DB 3212 AAGGCCCGGACGTCAGGACGTCGTGGCATCAATTGC 3247
RESULT 2
US-10-425-114-24930
; Sequence 24930, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 24930
; LENGTH: 3589
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3689-221-C10_FLI
US-10-425-114-24930
Alignment Scores:
Pred. No.: 0 Length: 3589
Score: 5520.00 Matches: 1030
Percent Similarity: 98.96% Conservative: 13
Best Local Similarity: 97.72% Mismatches: 7
Query Match: 97.73% Indels: 4
DB: 17 Gaps: 3
US-10-627-132-30 (1-1052) x US-10-425-114-24930 (1-3589)
QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
DB 247 ATGAGGCGCAGCGCGGCTTGTGGCGGCTCGCACACCGGACGAGCTGGTGTCTATC 306
QY 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGluIleCys 40
DB 307 CGGGGCCACAGGACCCCAAGCCGCTCGGCGGCTGAGCGGCGAGGTGTGCGAGATATGC 366
QY 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
DB 367 GCGCAGCAGGTGCGGCTCACGCTGAGCGCACCTTCTGCTGGCTGCAACGAGTGGCT 426
QY 61 PheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsnCysPro 80
DB 427 TTCCCGGCTGCTCCCGGCTTGTCTAGAGTACGAGCGCGGAGGGGACCGACGAACCTGCC 486
QY 81 GlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaAlaGlyAspAsp 100
DB 487 CAGTGAAGACGCGCTACAGCGCTCAAGGGGAGCGCGGCTCGCCGGGACGATGAC 546
QY 101 GluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnArg 120
DB 547 GAGGAGCAGATCGACGACCTGGAGCACGAGTTCAACATCGACGACGAG---AAGCAGAGG 603
QY 121 GlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMet 140
DB 604 CAGCTGGAGGGCAACATGACAGACCGCATCACCGGAGGCGCATGCTGCACGGCAAGATG 663

141 SerTyrGlyArgGlyProAspAspGlyAspGlyValAsnAsnThrProGlnIleProProile 160
Db AGCTACGGAGGGGGCGGACGACGGGAGGGCAACCAACCCCGAGATGCGCCCATC 723
161 IleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrGly 180
Db ATCACCGGCGCGCTCGTGGCGGTGAGCGGTGAGTTTCCAAATCACAAACGGGTATGGC 783
181 HisGlyGluValSerSerSerLeuHisValArgIleHisProTyrProValSerGluPro 200
Db CACGCGAGCTCTCGTCTCTCCACCAAGCGCATCCATCCCTGTTCTCGAGCCA 843
201 GlySerAlaTyrTrpAspGluLysGluValSerTrpLysGluArgMetAspAspTrp 220
Db GGGAGCGCAAGTGGGACGAGAGAGAGTAGTGTGAAGAGAGAGATGGACATGG 903
221 LysSerLysGlnGlyIleLeu-----GlyGlyGlyAlaAspProGluAspMetAspAla 238
Db AAGTCCAAGCAGGCGATCTCTGGCGGCGCGCGCGGACCCCGAAGACATGGACGCC 963
239 AspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSer 258
Db GACGTGCGCTGAAACGACGAGCGAGGCGCGCTGTTCGCGCAAGGTGTGATCGCGTCG 1023
259 SerLysValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPhePhe 278
Db AGCAAGGTGAACCCGTACCGATGTCATCGTGTGCGTCTCGTGTCTCGCTTCTTC 1083
279 LeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerIle 298
Db CTCGCTACCGATCTCTGACCCCTGTCGCGAGCGCATCGGCTGTGGCTGTCTCCATC 1143
299 IleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPhePro 318
Db ATCTCGAGATCTGGTTCCGCTGTCTGTGATCTCTGGACCATCGGCTGTGGCTGTCC 1203
319 IleAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGluPro 338
Db ATGACCGCGAGACGTACTGACCGCTCACCTTCAGGTACGAAGAGAGAGGAGCGG 1263
339 SerLeuLeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluProPro 358
Db TCGCTGTCTCGCGGTGACCTGTCTGTCAGCAGCGTGGACCCGCTCAAGAGCGCGCG 1323
359 LeuValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysVal 378
Db CTGGTCACCGCCAAACACCGTGTCTTCCATCTCTCGCGGTGGACTACCCCGTGGCAAGGTC 1383
379 SerCysTyrValSerAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThr 398
Db TCCTGCTACGCTCTCGACGAGCGCGCTGATGTCACGTTTCAGAGCCCTGTTCGAGAGC 1443
399 AlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArgAla 418
Db GCCGAGTTCCGCGGAGTGGGTGCGGTCTTCAGAGAAGTCTGCAATCGAGCCCGCGCC 1503
419 ProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThrPhe 438
Db CCGGAGTTCTACTTCTCGCTCAAGGTGATTAACCTCAAGGACAAGGTGCAGCCACCTTC 1563
439 ValGlnGluArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAla 458
Db GTGCGAGGAGCGCGCGCATGAAGAGAGATACGAGGAGTTCAAGTCCGATCAACCGG 1623
459 LeuValAlaLysAlaMetLysValProAlaGluGlyTyrIleMetLysAspGlyThrPro 478
Db CTGGTGGCCAAAGCCATGAAGTGTCCGCGAGAGGGGTGATCATGAAGGACGCGACGCG 1683
479 TrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSer 498
Db TGGCCCCGGGAACAACACCGCGACCACTTGGCATGATCCAGGTGTCTCTGGCGCCACAGC 1743

499 GlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLys 518
Db GCGCGCCACGACACCGGAGGCAACGAGTGTCCCGCTCGTGTACGTCTCCGTGAGAAG 1803
519 ArgProGlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSer 538
Db CGCCCGGCTTCCAGCACCAAGAGGCGCGCCATGAACGCTCTGATCCCGCTCTCC 1863
539 AlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsn 558
Db GCCGTGCTGACCAACCGCCCTTCATGCTCACTTGGATGTGATCACTACATCAACAC 1923
559 SerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysVal 578
Db AGCAAGGCCATCCGGGAGGCCATGTGCTTCTCATGACCTCAGGTGCGCGCAAGGTC 1983
579 CysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsn 598
Db TGCTACGTGCAAGTTCCTCAGAGGTTTCGACGGATCGACGTGCACGCCATACGCCAAC 2043
599 ArgAsnThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProVal 618
Db AGGAACACCGCTCTTCTTCGACATCAACATGAAGGGCTGGACGGCATCCAGGGCCCCGTC 2103
619 TyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProLys 638
Db TAGTCGGGACAGGTCGTGTTCCGCGCGCAGCGCTCTACGGCTACAACCCCTCCCAAG 2163
639 GlyProLysArgProLysMetValThrCysAspCysCysProCysPheGlyArgLysLys 658
Db GGGCCCAAGAGGCAAGATGGTGACCTGCGACTGTCTGCCCGCTCTCGCGCCGCAAGAAG 2223
659 ArgLysHisAlaLysAspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAsp 678
Db CGGAAGACCCCAAGGATGGGTGCGCGAGGGCACCCCGGATATC-----GTGGACAGCGAC 2280
679 LysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPhe 698
Db AAGGAGATGCTATGTTCCAGATGAACCTTCGAGAAGCGGTTCCGGCAGTCCCGCGCGTTC 2340
699 ValThrSerThrLeuMetGluGluGlyValProProSerSerSerProAlaAlaLeu 718
Db GTCACTCCAGCTCATGAGGAGGCGCGCTCCCTCGCTCGTACGAGCCAGCTGCGCTC 2400
719 LeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeu 738
Db CTCAGGAGGCCATCCACGTCATCAGCTGCGGTACGAGGACAGACGAGCTGGGGCTG 2460
739 GluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHis 758
Db GAGCTGGGGTGGATCTACGGGTCCATCACGAGGACATCTCTGACGGGTTCAAGATGACAC 2520
759 CysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAla 778
Db TGCCGCGGTGGCGCTCGGTGTACTGATGTCGGAAGCGCGCGCTTCAAGGGGTGCGCG 2580
779 ProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGlu 798
Db CCGATCAACCTGTCCGACCGTCTCAACAGGTGCTCCGGTGGGCGCTGGGGTTCGTCGAG 2640
799 IlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrp 818
Db ATCTTCTTACCGCGCACGCCCTCTGTGTACGGCTACAAGAACGCGCAACCTCAAGTGG 2700
819 LeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeu 838
Db CTGGAGCGCTTCGCTCATCAACACCACTATACCCCTTACCTCGCTCCCGCTGCTC 2760
839 AlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIle 858
Db GCCTACTGTACCTTGCACCGCGCTGTGCTCTCTACCGGCAGATTCATCATGCCGTGATC 2820
859 SerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIle 878

Db 2821 AGCAGTTGCGCGAGCTCTTCTTCATCGCCCTCTTCATGTCCATCTTCGCGACGGGCATC 2880
Qy 879 LeuGluMetArgTTPSerGlyValSerIleGluGluTTPTrpArgAsnGluGlnPheTTP 898
Db 2881 CTGGAGATGCGGTGAGCGGGGTGAGCATCGAGGAGTGGTGAGGACGAGCATCTTCGG 2940
Qy 899 ValIleGlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuValLeu 918
Db 2941 GTCATCGCGGGGTGCGCGCACCTCTTCGCGGTGCGTGCAGGGGCTGCTCAAGGTCTCTG 3000
Qy 919 AlaGlyIleAspThrAsnPheThrValThrSerIleAlaThrGlyAspGluAspGlu 938
Db 3001 GCCGGATGACACCACTTACCGTCACTTCAAGGCCACCGGGGAGCGTGAGCAGCAG 3060
Qy 939 PheAlaGluLeuTyrAlaPheIleTTPThrLeuLeuIleProGTPThrThrLeuLeu 958
Db 3061 TTGCGCGAGCTCTACGCGTTCAAGTGGACCGCTCTCATCCCGCCACCGCTGCTC 3120
Qy 959 IleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGln 978
Db 3121 ATCATCAACATCATCGCGGTGCGCGCGCATCTCCGACGCCATCAACAAACGGGTACCG 3180
Qy 979 SerTTPGlyProLeuPheGlyIleLeuPheAlaPheTTPValIleValIleTyr 998
Db 3181 TCGTGGGCGCGCTCTTCGCGAGCTCTTCGCGCTTCTGGGTCTCATCTCCACCTCTAC 3240
Qy 999 ProPheLeuIleGlyLeuMetGlyArgGlnAsnArgTTPProThrValValIleTTP 1018
Db 3241 CGTTCCTCAAGGGCTCATGGGCGCGCAGAACCGGACGCCCATCTGTCATCTGG 3300
Qy 1019 SerIleLeuLeuAlaSerIlePheSerLeuTTPValArgIleAspProPheIleVal 1038
Db 3301 TCCGTGCTGCTGGCTCCATCTCTCTCCCTGCTCTGGGTGAGTCCCTTTCATCTGTC 3360
Qy 1039 ArgThrIleGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3361 AGGACCAAGGGCGCGCGACGTGAGGAGTGTGGCATCAACTGC 3402

RESULT 3

US-10-437-963-15093/c
; Sequence 15093, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 15093
; LENGTH: 3448
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_20969C.1
US-10-437-963-15093

Alignment Scores:

Pred. No.:	0	Length:	3448
Score:	5347.50	Matches:	992
Percent Similarity:	96.78%	Conservative:	30
Best Local Similarity:	93.94%	Mismatches:	29
Query Match:	94.68%	Indels:	5
DB:	18	Gaps:	4

US-10-627-132-30 (1-1052) x US-10-437-963-15093 (1-3448)
Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
Db 3373 ATGAGGCGACGCGCGCGGTGGTGGCCGGGTGCGCAACCGGAACGAGTGGTGTGTATC 3314
Qy 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGluIleCys 40
Db 3313 CGGGGCGACGAGGAGCCCAAGCGCTGCGGGCGCTGAGCGGCGAGTGTGCGAGATATGC 3254
Qy 41 GlyAspGluValGlyLeuThrValAspClyAspLeuPheValAlaCysAsnGluCysGly 60
Db 3253 GCGCAGCAGGTGCGCGCGCACCGTTCAGCGCGACCTCTTCGTGCGCTGCAACGAGTGC 3194
Qy 61 PheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsnCysPro 80
Db 3193 TTCCCGGTGCGCGCGCTGTCTACGATACGAGCGCGCGCGGAGCACCGAGACTGCCCC 3134
Qy 81 GlnCysIleThrArgTyrLysArgLeuIleGlySerProArgValAlaGlyAspAsp 100
Db 3133 CAGTGAAGACCCGCTCAAGCGCTCAAGGGGAGCGCGGTGCGCGGCGAGCGAGGAC 3074
Qy 101 GluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnGln 119
Db 3073 GAGGAGGACATGTGAGCACCGTTCAGCATCTGACGACGAGCGAAGCAAGAGCAG 3014
Qy 120 --- ArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
Db 3013 CTGCAGCAGGATCAGGATGGCATGCGAAGACAGCACATCACCGAGCGCATGCTGCACGCG 2954
Qy 139 ArgMetSerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIlePro 158
Db 2953 AAGATGAGTACGGGAGGGGCGCGCGAGCGCGCGAGCGGCAACACACCGCGCGCG 2897
Qy 159 ProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGly 178
Db 2896 CCGATCATACCGCGCTCGCTCCGCTGAGCGGGGAGTTCGCCGATATCGAACACG 2837
Qy 179 TyrGlyHisGlyGluValSerSerLeuHisIleArgIleHisProTyrProValSer 198
Db 2836 CATGGCCATGGCGAGTTCCTCTCTCTGCGCAAGCGCATCCACCCCTACCCGCTGTCT 2777
Qy 199 GluProGlySerAlaLysTTPAspGlyLysGluValSerTTPLysGluArgMetAsp 218
Db 2776 GAGCAGGGAGTCAAGAGTGGGACGAGAAGAAAGAGGTGAGCTGGAAGGAGGATGAC 2717
Qy 219 AspTTPlysSerLysGlnGlyIleLeuGlyGlyAla---AspProGluAspMetAsp 237
Db 2716 GACTGGAATCAAGCAGGGCATCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2657
Qy 238 AlaAspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAla 257
Db 2656 GCCGAGTCCCATGACGACGAGCGAGGCGCGCGCTGTGAGGAAGGTGTGATTCGCG 2597
Qy 258 SerSerLysValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPhe 277
Db 2596 TCGAGCAAGGTGAACCGGTACCGGATGGTGTATCATCTCCGCTCTCGTGTGCTCGGCTTC 2537
Qy 278 PheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTTPLeuValSer 297
Db 2536 TTCTCTCGGTACCGCATCTCTCACCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 2477
Qy 298 IleIleCysGluIleTTPPheAlaIleSerTTPLeuAspGlnPheProLysTTP 317
Db 2476 ATCATCTGCGGAGTCTGGTTCGCGCGCTGTGCTGGATCTCTGACACGAGTCCCCCAAGTGGTAC 2417
Qy 318 ProfileAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGlu 337
Db 2416 CCGATCGACAGGAGACCTACCTCGACCGCTCTCCCTCCGCTACGAGCGCGAGGGGAG 2357
Qy 338 ProSerLeuLeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluPro 357

Db 2356 CCGTGGCTGTGTCGGGGTGGACCTGTTCTGTCAGCACGGTGGATCGCTCAAGGACGG 2297
Qy 358 ProLeuValThrAlaAenThrValLeuSerIleLeuAlaValAspTyrProValAspLys 377
Db 2296 CCGTGGTCAACGGCCAAACCGTGTCTCCATCTCCGCGTGCAGTACCCCGTGCACAAG 2237
Qy 378 ValSerCysTyrValSerAspAspGlyValSerMetLeuThrPheGluSerLeuSerGlu 397
Db 2236 GTGTCTGTGCTACGTCCTCCGACGCGCGGTCCATGTCTACGTTTCGAGTGCCTGTGCGAG 2177
Qy 398 ThrAlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArg 417
Db 2176 ACGGCGAGTTCGCGCCGCAAGTGGTGCCTTCTGCAAGAAGTTTCAGCATCGAGCCCGCG 2117
Qy 418 AlaProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThr 437
Db 2116 GCCCGGAGTCTACTTCTCCAGAAGTGCATACCTCAAGGACAAAGTCCATCCCAAC 2057
Qy 438 PheValGlnGluArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsn 457
Db 2056 TTCGTCCAGAGCGCCCGCCCATGAAGAGAGTACGAGAGTTCAAGGTGAGATCAAC 1997
Qy 458 AlaLeuValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThr 477
Db 1996 GCGTGGTGGCGAAGGCGCAGAAAGTCCGCGCGAAGGGTGGATCATGAAGGACGGGACG 1937
Qy 478 ProTrpProGlyAenAenThrArgAspHisProGlyMetIleGlnValPheLeuGlyHis 497
Db 1936 CCATGGCGGGGAACAACCCCGGACACCCCGGCGCATGATCCAGGTGTTCTCTCGGCCAC 1877
Qy 498 SerGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlu 517
Db 1876 AGCGGCGCCACGACACCGAGGGCAACGAGCTCCCGCCCTCGTCTAGCTCTCCGTGAG 1817
Qy 518 LysArgProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgVal 537
Db 1816 AAGCGCCCGGCTTCCAGCACCAACAAGAGCGCGCCCATGAACGCCCTCATTCGTGTG 1757
Qy 538 SerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsn 557
Db 1756 TCGGCGGTGCTGACGAACCGCCGTTCTATGCTCAACTTGGATTGCGATCATCATCAAC 1697
Qy 558 AsnSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLys 577
Db 1696 AACAGCAAGGCCATCAGGAGCGCATGTGCTCTCATGGATCCGAGGTCCGACGGAAG 1637
Qy 578 ValCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAla 597
Db 1636 GTTTGCTACGTGACGTTCCCGCAGAGGTTTCGACGCGCATCGAGTCCACGACCGATACGCC 1577
Qy 598 AsnArgAenThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyPro 617
Db 1576 AACCGCAACACCGTCTTCTGACATCAACATGAAGGGGCTTGATGGGATCCAGGGCCCG 1517
Qy 618 ValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProPro 637
Db 1516 GTGTACGTCCGACAGGGTGTGTTCAGGCGGCGAGGCGTGTACGGATACACCCACCC 1457
Qy 638 LysGlyProLysArgProLysMetValThrCysAspCysCysProCysPheGlyArgLys 657
Db 1456 AAGGACCAAGAGGCCCAAGATGTGACCTGCGACTGCTGCCCTTGTCTTCGGGAGGAAG 1397
Qy 658 LysArgLysHisAlaLysAspGlyLeuProGluGly----ThrAlaAspMetGlyValAsp 676
Db 1396 AAGCGGAAGCAGCGCAAGGACGGCTCCCGAGGCGCGTCCGCGCGCGAGGGATGGAC 1337
Qy 677 SerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAla 696
Db 1336 AGCGACAAGGATGCTCATGTCCGAGATGAACCTTCGAGAGCGGTTTCGGGACGTCCGG 1277
Qy 697 AlaPheValThrSerThrLeuMetGluGlyGlyValProProSerSerSerProAla 716
Db 1276 GCGTTCGTGACGTGACGTGATGAGGAAGGCGCGTCCCGCGTCTGTCAGCCCGCC 1217

Qy 717 AlaLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrp 736
Db 1216 GCGCTCTCAAGAGGCCATCCATGTCACTGCTGGGTACGAGGACCAAGACCGACTGG 1157
Qy 737 GlyLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLys 756
Db 1156 GGTCTCGAGCTGGGTGGATCTACGGTTCGATCAGCGAGGACATCTCAACGGGGTTCAAG 1097
Qy 757 MetHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaIlePheLysGly 776
Db 1096 ATGCACTGCGCGGTGGAGTCTGCTATCGCATCCGAAGAGGCGGCTTCAAGGGG 1037
Qy 777 SerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySer 796
Db 1036 TCAGCGCCGATCAACCTATCTGACCTTCAACAGAGTCTCCGGTGGGCGCTCGGCTCC 977
Qy 797 ValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeu 816
Db 976 GTCGAGATCTTCTTACGCGGCACAGCCGCTCTCTAGGCTTACAAGAACGGCAACCTC 917
Qy 817 LysTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuPro 836
Db 916 AAGTGGCTCGAGCGCTTCTCTATCAACACCACTTACCCCTTCACTTCTCTCCCC 857
Qy 837 LeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetPro 856
Db 856 CTCTCTGCGCTACGACCCCTTACCGCGTCTGCTCTCACCGGCAAGTTCATCATGCT 797
Qy 857 SerIleSerThrPheAlaSerLeuPheIleAlaLeuPheMetSerIlePheAlaThr 876
Db 796 CCGATTAGCACGTTTGGAGTTTGTCTTCTCATCGCTCTTCTCTCATCTTCTCGCGACG 737
Qy 877 GlyIleLeuGluMetArgTrpSerGlyValSerIleGluLeuTrpTrpArgAsnGln 896
Db 736 GGCATCTCGAGATGAGGTGGAGCGGTGAGCATCGAGAGTGGTGGAGAACGACGACG 677
Qy 897 PheTrpValIleGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLys 916
Db 676 TTCTGGGTCTATCGCGCGTGTGCGGCACCTGTTTGGGGTGGTGGAGGCGCTGCTCAAG 617
Qy 917 ValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAsp 936
Db 616 GTGTCGCGCGGATCGACACCACTTCCAGCTCAGCTCCAAGGCCACCGGAGACGAGGAC 557
Qy 937 AspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThr 956
Db 556 GACGAGTTCGCGAGCTTACGCTTCAAGTGGACCACTTCTTCTCGCTTCTCGGCTCCACCG 497
Qy 957 LeuLeuIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGly 976
Db 496 CTGCTCATCTCAACATCATCGCGCTGTCGCGCGCTTCCGAGCGCATCAACACGGC 437
Qy 977 TyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHis 996
Db 436 TCCGAGCGTGGGCGCGCTTTCGGGAAGCTTCTTCTCGCTTCTGGGTTCATCGTCCAC 377
Qy 997 LeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgTrpProThrValVal 1016
Db 376 CTCTACCCCTTCTCAAGGGGCTCATGGGAGGAGAACCGGACCGCCCATTTGTGTGTC 317
Qy 1017 IleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPhe 1036
Db 316 ATCTGTCGCTGCTGCTGCTGCTCATCTTTTCTTCTGCTGCTGCTGCTGCTGCTGCT 257
Qy 1037 IleValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 256 ACCATCAAGCGCAGGGCGCTTACGCTCAGGCGAGTGCAGCATCAACTGC 209

RESULT 4

US-10-059-059-29

; Sequence 29, Application US/10209059

; Publication No. US20030163838A1

; GENERAL INFORMATION:

; APPLICANT: Drugga, Kanwarpal S.
 ; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
 ; FILE REFERENCE: 0864R2
 ; CURRENT FILING DATE: 2002-07-31
 ; PRIOR APPLICATION NUMBER: 60/096,822
 ; PRIOR FILING DATE: 1998-08-17
 ; PRIOR APPLICATION NUMBER: 09/371,383
 ; PRIOR FILING DATE: 1999-08-06
 ; PRIOR APPLICATION NUMBER: 09/550,483
 ; PRIOR FILING DATE: 2000-04-14
 ; NUMBER OF SEQ ID NOS: 52
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 29
 ; LENGTH: 3028
 ; TYPE: DNA
 ; ORGANISM: Zea mays

US-10-209-059-29

Alignment Scores:

Pred. No.: 0 Length: 3028
 Score: 5055.00 Matches: 944
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 89.50% Indels: 0
 DB: 16 Gaps: 0

US-10-627-132-30 (1-1052) x US-10-209-059-29 (1-3028)

QY 109 HisGluPheAsnIleAspAspGluAsnGlnGlnArgGlnLeuGluGlyAsnMetGlnAsn 128
 Db 1 CACGAGTTCAACATCGACGACGAGAGATCAGCAGAGCGAGCTGGAGGGCAACATGCAGAAC 60
 QY 129 SerGlnIleThrGluAlaMetLeuHisGlyArgMetSerTyrGlyArgGlyProAspAsp 148
 Db 61 AGCCAGATCACCGAGGCGATGTCGACGCGCAGGATGAGCTACGGGAGGGGGCCCGCAGCAG 120
 QY 149 GlyAspGlyAsnAsnThrProGlnIleThrProIleIleThrGlySerArgSerValPro 168
 Db 121 GCGGACGGCAACACACCCCGCAGATCCGCCCATCATCACCGCTCCCGTCCGTCGCGG 180
 QY 169 ValSerGlyGluPheProIleThrAsnGlyTyrGlyHisGlyGluValSerSerSerLeu 188
 Db 181 GTGACGGGTGAGTTTCCGATTACCAACGGGTATGCCACGCGCGAGGTCTCGTCTTCCCTG 240
 QY 189 HisIysArgIleHisProTyrProValSerGluProGlySerAlaIlyTrpAspGluLys 208
 Db 241 CACAAGCGCATCCATCCGTACCTGTGTGAGCCAGGGAGTGCACAAAGTGGGACGAGAAAG 300
 QY 209 LysGluValSerTrpLysGluArgMetAspAspTrpLysSerLysGlnGlyIleLeuGly 228
 Db 301 AAAGAGTGAAGCTGGAGAGAGGATGACGAGCTGGAAGTCCAAAGCGGCATCTCCGCG 360
 QY 229 GlyGlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGln 248
 Db 361 GCGCGCGCGATCCCGAGACATGACCGCGAGCTGGCACTGAACGACGAGCGGAGGCGAG 420
 QY 249 ProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIle 268
 Db 421 CCGCTGTGAGAGAGGTGCGATCCGTCGTCGAGCAAGGTCAACCCGTACCGGATGGTGATC 480
 QY 269 ValValArgLeuValValIleAlaPhePheLeuArgTyrArgIleLeuHisProValPro 288
 Db 481 GTGGTGGCGTCTGTTGTGCTCGCTTCTTCCGTACCGGTATCCTGCACCCCGTCCCG 540
 QY 289 AspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrp 308
 Db 541 GACGCCATCGGGCTGTGCTCTCCATCATCTCGAGAGATCTGGTTCGCCCATCTCTCTGG 600
 QY 309 IleLeuAspGlnPheProIlyTrpPheProIleAspArgGluThrTyrLeuAspArgLeu 328

Db 601 ATCTCTCGACCAAGTTCCCAAGTGGTTCCTCCATCCAGCCGAGAGCTACTCGACCGCTC 660
 QY 329 SerLeuArgTyrGluArgGluGlyProSerLeuLeuSerAlaValAspLeuPheVal 348
 Db 661 TCCCTCAGGTACGAGAGGAGGAGCGCTCGTCTGTCTGCGCGGTGAGACTGTTCGTG 720
 QY 349 SerThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIle 368
 Db 721 AGCACGGTGGACCCGCTCAGGAGCCGCGCTGGTGTGACCGCAACACCGTGTCTTCCATC 780
 QY 369 LeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSer 388
 Db 781 CTCGCGGTAGACTACCCCGTGGACAGGTCTCTGCTACGTCTCCGACGAGCGCGCGTGG 840
 QY 389 MetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPhe 408
 Db 841 ATGCTGAGCTTCGAGTCCGCTCGGAGACGCGGAGTTCGCGCGCAAGTGGTGGTCTTC 900
 QY 409 CysLysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAsp 428
 Db 901 TGCAGAAGTTCGGCATCGAGCCCGCGCCCGGAGTTCTACTTCTCGCTCAAGGTCGAC 960
 QY 429 TyrLeuLysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArgGlu 448
 Db 961 TACCTCAAGGACAAAGGTGACGCCACCTTCTGTGAGGAGCGCGCGCATGAAGAGAGAG 1020
 QY 449 TyrGluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAla 468
 Db 1021 TATGAGGAGTTCAGGTCCGANTCAACGGCTGTGTGGCCAGGCGCATGAAGTGGCGCA 1080
 QY 469 GluGlyTrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisPro 488
 Db 1081 GAGGGTGGATCATAGGACGCGACGCGGTGGCCCGGGAACAACACCGCGGACCCACCCC 1140
 QY 489 GlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeu 508
 Db 1141 GGCATGATCCAGGTGTCTCTGGGCCACAGCGCGCGCCACGACCGAGGGCAACAGAGTGT 1200
 QY 509 ProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAla 528
 Db 1201 CCCGCCCTCGTGTAGTCTCCCGTGAGNAGCGCCCGGGATTCCAGCACCAACAGAGGCC 1260
 QY 529 GlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeu 548
 Db 1261 GCGCCATGAACGCTCTGATTCCGGTCTCCCGCGTGTGACCAACCGCGCCATTCTATGCTC 1320
 QY 549 AsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPhe 568
 Db 1321 AACTTGGACTGTGATCACTACATCAACAACAGCAAGGCCATCCGGGAGGCCATGTGCTTC 1380
 QY 569 LeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAsp 588
 Db 1381 CTCTATGACCTCAGTTCGCGCGAGAGGTCTGCTACGTTCAGTTCGCCGAGAGTTCGAC 1440
 QY 589 GlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMet 608
 Db 1441 GGCATCGAGCTGCACGACCGCATACGTAACAGGAACACCGTCTTCTTCGATCAACATG 1500
 QY 609 LysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArg 628
 Db 1501 AAGGGCTGGACGGCATCCAAAGGCCGCTGTACGTGCGGACAGGGTGTGCTTCCGGGCG 1560
 QY 629 GlnAlaLeuTyrGlyTyrAsnProLysGlyProLysArgProLysMetValThrCys 648
 Db 1561 CAGCGCTCTACGGCTACACCTCCCAAGGAGCCCAAGAGGCCCAAGATGGTGACCTGC 1620
 QY 649 AspCysCysProCysPheGlyArgLysLysArgLysHisAlaLysAspGlyLeuProGlu 668
 Db 1621 GACTGCTGCCGTGCTTCCGGCGGCAAGAGCGGAAACACGCGCAAGGACGGGTGCGGAG 1680
 QY 669 GlyThrAlaAspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPhe 688

1681	Db	GGCACCCTGATATGGGAGTAGATAGCGACAAGGAGATGCTCATGTGCCACATGAACCTTC	1740
689	Qy	GLuLySerGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGluGlyGly	708
1741	Db	GAGNAGCGGTTGGGCAGTCCGCGCGTTCGTCACTCGACGCTGATGGAGGAGCGGC	1800
709	Qy	ValProProSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCys	728
1801	Db	GTCCCTCTTTCGTAGACCCCGCGCTCTCTCAAGGAGGCCATCCATGTTCATCAGCTGC	1860
729	Qy	GlyTyrGluAspLysThrAspTrpGlyLeuGluLeuGlyTrpIleTyrGlySerIleThr	748
1861	Db	GGCTACGAGGACCAAGACCGACTGGGGCTGGAGCTGGGGTGGATCTACGGGTGGATCAG	1920
749	Qy	GluAspIleLeuThrGlyPheLysMetHisCysArgGlyTrpArgSerValTyrCysMet	768
1921	Db	GAGGACATCTGACGGGTTCAAGATGCATCTGCCGCGGGTGGCGCTCCGTGTACTGCATG	1980
769	Qy	ProLysArgAlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGln	788
1981	Db	CCGAAGCGGCGGGTTCAGGGGTGGCGCGCATCAATCTACGGACCGCTCTCAACCAG	2040
789	Qy	ValLeuArgTrpAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeu	808
2041	Db	GTGCTCGGTGGGCGCTGGGTCCGTCGAGATCTTCTTCAGCCGGCACACGCCCTGCTG	2100
809	Qy	TyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrThr	828
2101	Db	TACGGCTACAGAAGACGGCACTCTAAGTGGCTGGAGCGCTTCGGCTACATCAACACC	2160
829	Qy	IleTyrProPheThrSerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeu	848
2161	Db	ATCTACCCCTTCACTCGCTCCGCTGCTCGCTTACTGCACCTCCCGCGCTCTGCCTC	2220
849	Qy	LeuThrGlyLysPheIleMetProSerIleSerThrPheAlaSerLeuPheIleAla	868
2221	Db	CTCACCGGCAGATTTCATCATGCCCTCGATTAGCACGTTCCGCCAGCTCTTCTTCATCG	2280
869	Qy	LeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIle	888
2281	Db	CTCTTCATGTCCATCTTCGCGACGGGCATCTCTGGAGATGCGTGGAGCGGGTGAGCATC	2340
889	Qy	GluGluTrpTrpArgAsnGluGlnPheTrpValIleGlyGlyValSerAlaHisLeuPhe	908
2341	Db	GAGGAGTGGTGAGGAAACGACGAGTCTCTGGGTTCATCGCGCGGTCTCCGCGCATCTCT	2400
909	Qy	AlaValValGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThr	928
2401	Db	GCCCTCGTCGACGGGCTGCTCAAGGTCTCTCGCCGGGATCGACACCAACTTCACCGTCA	2460
929	Qy	SerLysAlaThrGlyAspGluAspGluPheAlaGluLeuTyrAlaPheLysTrpThr	948
2461	Db	TCCAAGGCCACCGCGACGAGGACGAGTTCGCCGAGCTCTACGGCTTCAAGTGGAC	2520
949	Qy	ThrLeuLeuIleProProThrThrLeuLeuIleIleAsnValIleGlyValValAlaGly	968
2521	Db	ACGCTCTCTCATCCGCCCCACACACGCTGCTCATCTAATTAACGTCTACGGGTCTG	2580
969	Qy	IleSerAspAlaIleAsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhe	988
2581	Db	ATCTCCGACGCCATCAACAACGGGTACCAGTCTCTGGGGGCCCTCTTCGGCAAGCTCT	2640
989	Qy	PheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGln	1008
2641	Db	TTCGCCCTCTCGGTTCATCGTCCACCTCTACCCGTTCCTCAAGGGGCTCATGGGCGC	2700
1009	Qy	AsnArgThrProThrValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeu	1028
2701	Db	AACAGGACCCCAACGGTTGTGTTCATCTGGTCCATCTCTGCTGGCTCCATCTTCTCC	2760
1029	Qy	LeuTrpValArgIleAspProPheIleValArgThrLysGlyProAspValArgGlnCys	1048
2761	Db	CTCTGGGTTCAGATCGACCTTTTCATCGTCAGGACCAAGGCCCGGACGTCAGGACGT	2820

Qy 1049 GlyIleAsnCys 1052
Db 2821 GGCATCAATTGC 2832

RESULT 5

```

US-09-900-237-13
; Sequence 13, Application US/09900237
; Patent No. US200201201241
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 3517
; TYPE: DNA
; ORGANISM: Glycine max
US-09-900-237-13

```

608 ATTGCTGGTGGTCTTCTAGGCGCTGTAAGTGGGAGTTCCCAATATCATCTAATGCTTAT 667
180 GlyHisGlyGluValSerSerSerLeuHisLysArgIleHisProTyrProValSerGlu 199
668 GGGGATCAGATGTTATCTCTCTTCTACCTGCATATAAGAGTGCATCCATCTCACTGCTGAA 727
200 ProGlySerAlaLysTrpAspGlu-LysLysGluValSerTrpLysGluArgMetAspAs 219
728 CTGGAAGTGCAGAGTGGACGAAAGGAAAAAGAGATGATGGAAGATAGAGATGATGA 787
219 pTrpLysSerLysGlnGlyIleLeuGlyGlyGlyAlaAspProGlu---AspMetAspAl 238
788 CTGGAATTCGACGAGCAATTTGGG-----CCTGAACCGGATGAGATCC 835
238 aspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSe 258
836 AGATGCGAGCATGTTAGATGAGCAAGGCAACCACTGTCAGAGAAAGTGCATAGCATC 895
258 rSerLysValAsnProTyrArgMetValIleValArgLeuValValLeuAlaPhePh 278
896 CAGCAAAATCAATCCATATAGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 955
278 eleuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerI 298
956 CTGAGATACAGACTCATGAAACCCAGTACATGATGCCCTGGGCTATGGCTAACTCTAT 1015
298 eileCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPhePr 318
1016 CATATGTGAATCTGTTTCTGTTTCTGATGATCTGGATCAGTTTCCCAATGGTTTCC 1075
318 oileAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyPr 338
1076 CATTTAGAGAGACTACTCTGACCGCTTTTCCATCAGGTATGAGCGTGAAGGTGAACC 1135
338 oSerLeuLeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluProPr 358
1136 AAACATGCTGCTCTAGATAGTGTGTTTGTAGTACCGTGGATCCCATGAAGAACCTCC 1195
358 oleuValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysVa 378
1196 TCTGTTACAGCAACACATGTTCTTCAATCTTGGCCATGGATTACCCGGTTGATAAAT 1255
378 lSerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluTh 398
1256 ATCATGCTACATTTCTGATGATGAGGAGCTCAATGTGTACATTTGAGTCTTATCAGAAC 1315
398 rAlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArgAl 418
1316 TGCAGATTTGCTAGAAAGTGGTACCGTTTGTAAAGAAATTTTCCATAGAACCTCGGGC 1375
418 aProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThrPh 438
1376 ACCTGAGATGATCTTTCAGCGAAGATGACTACCTAAAGGACAAGTGCACACCTT 1435
438 eValGlnGluArgArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAl 458
1436 TGTTAAGGAGCGTCGAGCTATCAAGAGGGAATACCAAGAGTTTAAAGTTAGGATCAATGC 1495
458 aleuValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrPr 478
1496 ACTTGTGCTAAGGCCCGAAAGTTCTTCAGGAGGATGGATCATGAGGATGGGACACC 1555
478 oTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSe 498
1556 ATGCCAGGGAATAACATAAGGATCATCTGATGATGATCAAGTGTCTTCTGGTAGCAG 1615
498 rGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLy 518
1616 TGGAGGCTTGATCTAGAGGAAACCAACTTCTCGCTGTTTATGTTTATGTTTCCAGAGAA 1675
518 sArgProGlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSe 538
1676 AAGGCTGTTTTCACCAACCAAGAAAGCTGGTGCCATGATGCTCTGTTTGGGTGATC 1735

538 rAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAs 558
1736 TGCTGTTCTCACAATAATGCTCTCTTTCATGTTGAACCTGGATTTGTGATCACTATGCTCAATAA 1795
558 nSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysVa 578
1796 CAGCAAGGCTGCCAGAGGCGCATGTTCTTGTATGGACCCACAAACTGGGAAGAAGGT 1855
578 lCysTyrValGlnPheProGlnAtqPheAspGlyIleAspValHisAspAtqTyrAlaAs 598
1856 CTGCTATGTCAGATTTCTCTCAAGATTTGATGTTATGATACACATGATCGTTATGCCAA 1915
598 nArgAsnThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProVa 618
1916 CAGGAACACAGTTTCTTCTTATTAATTAACATGAAGGCTCTAGATGTTTCAAGGCTCTGT 1975
618 lTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProProLy 638
1976 ATATGTTGGGACTGGATGTTTTCAGGAGCAAGCTTTGTATGGCTATAATCTCTCCAA 2035
638 sGlyProLysArgProLysMetValThrCysAspCysCysProCysPheGly---ArgLy 657
2036 GGGTCCAAGGCTGCCAAAATGGTAAGCTGTGATTTGGCCGCTGTTTGGAGCGCGCA 2095
657 sLysArgLysHisAlaLysAspGlyLeuProGluGlyThrAlaAspMetGlyValAspSe 677
2096 GAAGTATTAAGGAGAAGATGATGCAATGGAGAGGCTCAAGCCTAAAGGAGTGAATGA 2155
677 rAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAl 697
2156 TGACAAAGAGGTGTGATGTCCTCCAAATGAATTTTGAAGAAGAAATTTGACAACTCTCTAT 2215
697 apheValThrSerThrLeuMetGluGluGlyValProProSerSerSerProAlaAl 717
2216 TTTTGTGACTTCTACCTTGATGGAAGGTTGGTGGCTCTCTCTTCAAGTCCAGGTGC 2275
717 aLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGl 737
2276 CTGCTTAAAGAGCAATTCATGTTAGCTGTGATGATGAAGATATAAAGTGAATGGG 2335
737 yLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMe 757
2336 ACTTGAGCTTGGTGGATCTATGGATCTATCAGAGATATTTCAACAGGTTTAAAGAT 2395
757 tHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySe 777
2396 GCATTTGGCTGGTGGAGTCCATTTATGTTATGCCAAGAGAGCTGCTTCAAGGGTAC 2455
777 rAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerVa 797
2456 TGCTCTCTATCAACTTGTCCAGATCGTCTCAACAGGTTCTTCTGTTGGGCACTTGGTTCCAT 2515
797 lGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLy 817
2516 TGAGATTTCTTTTATGTCACCTGCTCTATGGTATGGTTCAGGAGAAAGAACTAAA 2575
817 sTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProle 837
2576 GTGGCTTGAGAGATTTGGCTATGCAACACACATGCTCTATTCATTCACCTCTCTCTCT 2635
837 uLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSe 857
2636 AGTTGCTCTACTGTTCTTCTCCAGCAGTTTGTCTACTCACTGCAAAATTCATCATGCCACC 2695
857 rIleSerThrPheAlaSerLeuPheIleAlaLeuPheMetSerIlePheAlaThrGl 877
2696 GATTAGCACCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2755
877 yIleLeuLeuMetArgTrpSerGlyValSerIleGluLeuTrpTrpArgAsnGluGlnPh 897
2756 TATTCTTGAGTTGAATGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 2815

Qy 897 eTTPValIleGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuValVa 917
Db 2816 TTGGGTCAATTGGTGTATCAGCTCACCTCTTTTGGTTATATAAAGGTCTGCTAAAGGT 2875
Qy 917 lLeuAlaGlyIleAspThrAnPheThrValThrSerLysAlaThrGlyAspGluAspAs 937
Db 2876 TCTGGCTGGAAATTGACACCAATTTCACTGTTTACATCAAGGCAACA-----GATGATGA 2929
Qy 937 pGluPheAlaGluLeuTyrAlaPheLysTTPThrThrLeuLeuIleProProThrThrLe 957
Db 2930 AGAGTTTGGAGAAATTGACACCTTTAAGTGAGCTACACTCTTGATTCCTCAACCACTAT 2989
Qy 957 uLeuIleAlaValIleGlyValValAlaGlyIleSerAspAlaIleAsnAnGlyTy 977
Db 2990 TTTGATCATTAACATTGTTGGTGTGTTGCTGGAATCTCAGATGCCATAAACAATGGTA 3049
Qy 977 rGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLe 997
Db 3050 CCAATCTCGGGACCACTCTTTGGAAGCTCTTCTTTCTCTGCGGTGATTTGCCATCT 3109
Qy 997 uTyrProPheLeuLysGlyLeuMetGlyAtqGlnAsnArgThrProThrValValIle 1017
Db 3110 CTATCCATTCCTTAAAGGTTTGAATGGTTCGCCAAATCGCACCCACCACTGTTGTGAT 3169
Qy 1017 eTrpSerIleLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspPropheIl 1037
Db 3170 TTGGTCAGTGCTATTGGCTCTATTCTTCTCTTACTTTGGGTAAGAATTGATCCATTGT 3229
Qy 1037 eValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3230 CCTCAAGACTAAGGACCTGATACCAAGCTATGTGAATCAACTGC 3275

RESULT 6

US-10-425-115-139789
; Sequence 139789, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 139789
; LENGTH: 2830
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_58974C.1
US-10-425-115-139789

Alignment Scores:
Pred. No.: 0 Length: 2830
Score: 4561.00 Matches: 856
Percent Similarity: 99.08% Conservative: 8
Best Local Similarity: 98.17% Mismatches: 6
Query Match: 80.75% Indels: 3
DB: 18 Gaps: 1

US-10-627-132-30 (1-1052) x US-10-425-115-139789 (1-2830)

Qy 183 GluValSerSerLeuHisLysArgIleHisProTyrProValSerGluProGlySer 202
Db 2 GAGCTCTCGTCTTCCCTGCACAAAGGCGCATCCATCCCTACCTCTGTCTGAGCGAGGAGC 61
Qy 203 AlalysTrpAspGluLysGluValSerTrpLysGluArgMetAspTrpLysSer 222
Db 62 GCCAAGTGGGAGGAGAAAGAGTG-AGCTGAAGGAGAGGATGACGACTGGGAAGTCC 120

Qy 223 LysGlnGlyIleLeu-----GlyGlyGlyAlaAspProGluAspMetAspAlaAspVal 240
Db 121 AAGAGGGGCATCTCG 180
Qy 241 AlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSerSerLys 260
Db 181 CCCTGAACAGACAGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Qy 261 ValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPhePheLeuArg 280
Db 241 GTGACCCCGTACCGGATGGTATCGTGGTGGCTCGTGGTGGCTCGTGGTGGCTCGTGG 300
Qy 281 TyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerIleLeuCys 300
Db 301 TACCGCATCTCGACCTGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Qy 301 GluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsp 320
Db 361 GAGATCTGGTTCGCGCGCTCTCGATCTCTGGATCTCTGGACGAGTTCCCAAGTGGTTCC 420
Qy 321 ArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGluProSerLeu 340
Db 421 CGCGAGAGCTACTGGACCCCTCACCTCAGGTACGAGAGGAGGAGGAGGAGGAGGAGG 480
Qy 341 LeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluProProLeuVal 360
Db 481 CTGTGCGCGGTGGACCTGTTGCTCAGCAGCGGTGGACCCGCTCAAGGAGCGCGCGCGGT 540
Qy 361 ThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCys 380
Db 541 ACCGCCAACACCGGTGCTCTCATCTCTCGCGGTGGACTACCCCGTGGACAAGGTCTCT 600
Qy 381 TyrValSerAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGlu 400
Db 601 TAGTCTCCAGCAGCGCGCGGTGATGCTCAGCTTCGAGCGCGGTGTCGGAGCGCGCG 660
Qy 401 PheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArgAlaProGlu 420
Db 661 TTGCGCGGGAAGTGGGTGGCTCTCTGCAAGAAGTTCTGTCATCGAGCGCGCGCGCGCG 720
Qy 421 PheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThrPheValGln 440
Db 721 TTTCTACTTCTCGTCAAGGTGATTTACTCAAGGACAAGGTGACGCCACCTTCGTGCG 780
Qy 441 GluArgArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAlaLeuVal 460
Db 781 GAGCGCGCGCGCATGAAGAGAGTACAGAGGTTCAAGGTTCCGATCCGATCAACGCGCT 840
Qy 461 AlalysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrProTrpPro 480
Db 841 GCCAAGGCCATGAAGGTGCGCGCGAGGGGTGGATCATGAAGGACGCGCACGCCGTGG 900
Qy 481 GlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGly 500
Db 901 GGGAAACAACACCGCGCGCACCCCTCGGCATGATCCAGGTGTTCTGGGCGCACAGCGG 960
Qy 501 HisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgPro 520
Db 961 CACGACACCGAGGCGCAAGAGCTGCCCGCGCTCGTGTACGTTCTCCGCTGAGAGCGCG 1020
Qy 521 GlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSerAlaVal 540
Db 1021 GGCTTCCAGCACCAACAAGGCGCGCGCATGAACGCTCTGATCCGCGCTCTCGCGCGGT 1080
Qy 541 LeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLys 560
Db 1081 CTGACCAACGCGCGCTTCTGCTCACTTGGACTGTGATCTACTACTACTCAACAAAGCAG 1140
Qy 561 AlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysValCysTyr 580
Db 1141 GCCATCCGGGAGGCGCATGTGTTCTCTCATGACCCCTCAGGTTCGCGCGGAGGTCTG 1200
Qy 581 ValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsnArgAsn 600

Db 1201 GTGCAGTTCCTCAGAGGTTCCAGCGCATCGAGCTGCACGACCGATACCGGACAGGAAC 1260
QY ThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProValTyrVal 620
Db 1261 ACCGCTTCTTCCACATCAACATGAAGGGGCTGGACGGCATCCAGGGCCCGGTCTACGTC 1320
QY GlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProProlLysGlyPro 640
Db 1321 GGGACAGGGTGGGTGTTCCGGGGCCAGGGGCTCTACGGGTACAAACCTCCCAAGGGCCCC 1380
QY LysArgProLysMetValThrCysAspCysCysProCysPheGlyValArgLysLysArgLys 660
Db 1381 AGAGGCCAAAGATGGTACCTGGACTGCTGCGGTGTTTCGGCGCGAAGAGCGGAG 1440
QY HisAlaLysAspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAspLysGlu 680
Db 1441 GACGCCAAGGATGGGCTGGCGGAGGACCGCGGATATCGAGTGGACGACGACGAGGAG 1500
QY MetLeuMetSerHisMetAsnPhenGluLysArgPheGlyGlnSerAlaAlaPheValThr 700
Db 1501 ATGCTCATGTCCTCAGATGAACATTCGAGAAAGCGGTTTCGGGCGAGTTCGGCGGCTTCGTCAG 1560
QY SerThrLeuMetGluGluGlyValProProSerSerSerProAlaAlaLeuLeuLys 720
Db 1561 TCCAGCTCATGGAGGAGGCGGCTCCTCGTGTGAGGCCAGCTCGCTCTCTCAAG 1620
QY GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeu 740
Db 1621 GAGGCCATCCAGCTCATCAGCTGCGGTACGAGGACCAAGCCAGCTAGGGGCTGGAGCTG 1680
QY GlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg 760
Db 1681 GGGTGGATCTACGGGTCCATCAGGAGGACATCTTGACGGGTTCGAAGTGCACCTGCCG 1740
QY GlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaPheLeu 780
Db 1741 GGGTGGGCTCGGTGTACTGCTGATGCGAAGCGGGCGGCGTTTCAGGGGTTCGGCGCGATC 1800
QY AsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePhe 800
Db 1801 AACCTGTCCGACCGTCTCAACAGGTGCTCGGTGGGCTGGGCTCGGTGAGATCTTC 1860
QY PheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGlu 820
Db 1861 TTCACCGCGCACAGCCCCCTGTCTACGGCTCAAGAACCGCAACCTCAAGTGGCTGGAG 1920
QY ArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyr 840
Db 1921 CGCTTCGCTCATCAACACCAACCACTACCCCTTCACCTCGCTCCGCTGCTCGCCTAC 1980
QY CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr 860
Db 1981 TGTAACCTGCCAGCGTCTGCTCTCCACCGCAAGTTTCATCATCGCGTTCGATCAGCAG 2040
QY PheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGlu 880
Db 2041 TTCGCAGCGCTCTCTTCATCGCGCTTTCATGTCCATCTTCGCGACGGGCATCTCGGAG 2100
QY MetArgTrpSerGlyValSerIleGluGluTyrTrpArgAsnGluGlnPheTrpValIle 900
Db 2101 ATGCGGTGAGCGGGGTGAGCATCGAGGAGTGGTGAGAAACGAGAGTTCGTGGGTGATC 2160
QY GlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAlaGly 920
Db 2161 GCGCGCGGTTCGCGCATCTCTTCGCGCTCGTGCGAGGGCTGCTCAAGGTCTCTGGCGGG 2220
QY IleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPheAla 940
Db 2221 ATCGACCAACATTCACCGTCAATCCCAAGGCCACCGCGCAGCAGGAGTTCGCGC 2280
QY GluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeuLeuIle 960

Db 2281 GAGCTCTACGCGTTCAAGTGGACCGCTCTCTCATCCCCCCACACCGCTGCTCATCATC 2340
QY AenValIleGlyValValAlaGlyLysSerAspAlaIleAsnAsnGlyTyrGlnSerTrp 980
Db 2341 AACATCATCGCGTCTGGCGCGCATCTCCGACGGCATCAACAACGGGTACCACTCGTGG 2400
QY GlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPhe 1000
Db 2401 GCGCCGCGCTCTTCGGCAAGCTCTTCGCGCTTCTGGGTCTCATCCACTCTACCGCTTC 2460
QY LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValValIleTrpSerIle 1020
Db 2461 CTCAGGGGCTCATGGGCGCCAGAACCGACGCCCATCTGTCTCATCTGGTCCGTG 2520
QY LeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThr 1040
Db 2521 CTGCTGGCTCTCATCTCTCCCTGCTCTGGGTGAGATCGACCCCTTTCATCGTCAGGACC 2580
QY LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 2581 AAGGGGCGCGAGCTCAGGCGAGTGTGGCATCAACTGC 2616

RESULT 7

US-10-393-840-69
; Sequence 69, Application US/10393840
; Publication No. US20030229922A1
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; TITLE OF INVENTION: Materials and Methods for the
; FILE OF INVENTION: Modification of Plant Cell Wall Polysaccharides
; FILE REFERENCE: 11000.1012c3
; CURRENT APPLICATION NUMBER: US/10/393,840
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 09/636,800
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 09/170,862
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/148,426
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT NZ/99/00169
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 956
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 3851
; TYPE: DNA
; ORGANISM: Pinus radiata
US-10-393-840-69

Alignment Scores:
Pred. No.: 0 Length: 3851
Score: 4477.00 Matches: 839
Percent Similarity: 84.56% Conservative: 92
Best Local Similarity: 76.20% Mismatches: 104
Query Match: 79.27% Indels: 66
DB: 17 Gaps: 16

US-10-627-132-30 (1-1052) x US-10-393-840-69 (1-3851)

QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValIle 20
Db 153 ATGGAAGCCAGCGCGGCTTGTTCGCCGTTCTCTATAACAGAAACGAGTTCGTGGTCAATC 212
QY 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGluIleCys 40
Db 213 CATGACATGAGGAGCCGAGCCCTTGAACACGTTGAGTGGCCAGCTCTGCCAGATTGT 272
QY 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
Db 273 GCGAGGACGTCGGGCTTAAACACAGCGGAGCTGTTCGTTCCTGTAATGAGTGGGG 332
QY 61 PheProValCysArgProCysTyrGluArgArgGluGlyThrGlnAsnCysPro 80

Db 333 TTTCTGTGTGTGGCGGTGTATAGTACGAGACGAGAGGAAATCAAGTCGTGGCCG 392
Qy 81 GlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAspAsp 100
Db 393 CAGTGCATACTCTTACAGCGTCAAAAAAGGAGTCCACGGGTGGAGGTGACGATGAT 452
Qy 101 GluGluAspLeuAspLeuGluHisGluPheAenIleAspAspGluAenGlnArg 120
Db 453 GAAGAAGACGTGTATGACATAGACATGATTTATGTGGAGATCAGCAAGAAACAGG 512
Qy 121 GlnLeuGluGlyAenMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMet 140
Db 513 CAG-----CAGATCACCGAGGCGATGCTCCACGAGCGCATG 548
Qy 141 SerTyrGlyArgGlyProAspAspGlyAspGlyAsn-----AsnThrProGlnIlePro 158
Db 549 AGCTATGGCGAGGTCCGACGACGAAAAATTCGCAGATTTGCTCATATATCCAGAGCTTCCT 608
Qy 159 -----ProIleIleThrGlySerArgSerValProValSerGlyGluPheProIle 175
Db 609 CCGCAGATTCCTGTACTTGCACAAACGGCCACTCGGT---GTGAGTGGGAGATTCCA-- 662
Qy 176 ThrAsnGlyTyrGlyHisGlyGluVal-----SerSerSerLeuHisLysArgIleHis 193
Db 663 ACGTCACTACTACGACGACAAACCAATTTGCTGCCAACCCCTGCAATGCTGAAGCGTGTGCAT 722
Qy 194 ProTyrProValSerGluProGlySerAlaLys-----TrpAspGluLysGlu--- 210
Db 723 -----CCAACTCCGAGCCGGGAGTGAAGGATCATCATGATCCAAACAGGAGATTT 776
Qy 211 -----ValSerTrpLysGluArgMetAspAspTrpLysSerLys 223
Db 777 GGTCTTATGCTTTGGGAACTGTCTTGGAGGAGCGGCGATGGTTATAAATCGAG 836
Qy 224 GlnGlyIleLeuGly-----GlyGly--- 230
Db 837 GAAAAAATAACAGGCCCATGATGATACGGAAGGAGATATCAATATAATGGGGGTTT 896
Qy 231 -----AlaAspProGluAsp---MetAspAlaAspValAlaLeuAenAspGluAlaArg 247
Db 897 GCACCAATAGCCCTGAAGATATATATGATCCCGATATGCCAATGACCGATGAAGCAAGG 956
Qy 248 GlnProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetVal 267
Db 957 CAGCCACTGTCCGAAAAGTGCCATTCCTTCACGACAAATAAATCCATACCGAATGGTC 1016
Qy 268 IleValValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProVal 287
Db 1017 ATTTGTAATTCGACTGATAGTGTGGGTATTTTCTCCGCTATCGTCTCTGAAATCCAGTG 1076
Qy 288 ProAspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSer 307
Db 1077 AAGAAATGATATATGGGCTCTGGGCCACTTCTATCGTTTGTGAATCTGGTTTGGCTTGTCA 1136
Qy 308 TrpIleLeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrIleuAspArg 327
Db 1137 TGGATCTTGATCAGTTTCCCAAGTGTGTGCTATCAGTCGTGGAACGTAATCTTGATCGA 1196
Qy 328 LeuSerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPhe 347
Db 1197 CTGTCAATTAAGGTACGACGAGAGGCGAACCATCAATGCTTGGCACCTGTGTGACCTCTTT 1256
Qy 348 ValSerThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSer 367
Db 1257 GTGAGTACTGTAGATCCACTGAAGAGAGCCCTCTTTGGTTACTTGCCAAATACAGTATTATCA 1316
Qy 368 IleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAla 387
Db 1317 ATCTCTTCAGTAGACTACCTCTAGACAAATGTCTCTGTATGTCTCTGATGACGAGCG 1376
Qy 388 SerMetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValPro 407
Db 1377 TCGATGCTTACTTTTGAATCTCTCTCGAGACCTCAGAAATTTGCCAGAAAATGGGTACCA 1436

Qy 408 PheCysLysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysVal 427
Db 1437 TTCTGCAAGAAATTCGACATTTGAGCCTCGCGCTCCCGAAATCTATTCTCTCGAATAATT 1496
Qy 428 AspTyrIleuLysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArg 447
Db 1497 GACTATCTGAAGGACAAATTTCAACCCACCTTTGTCAAAGAGCGCGCTGCATGAAGAGA 1556
Qy 448 GluTyrGluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValPro 467
Db 1557 GAATATGAAGAAATTCAGGTGCGCATCAATCGGTGGTTGGTTCGAAAGGCCCTCTAAAGTGCCT 1616
Qy 468 AlaGluGlyTrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHis 487
Db 1617 AAGNAGGATGGCAATGCAAGACGGTACGCTTGGCTTGGTAAATAATACCCGTGACCAT 1676
Qy 488 ProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGluAsnGlu 507
Db 1677 CTTGGTATGATCCAAAGTGTCTTTGGGTACAGTGGCGGCTCGATACAGAGGCAATGAG 1736
Qy 508 LeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLys 527
Db 1737 CTTCTCGGCTAGTATATGTTTCTCGGAGAGAGACTGTTTCCAGCATCACAGAGAG 1796
Qy 528 AlaGlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMet 547
Db 1797 GCCGTGCCATGAATGCTTTGGTTCGGGTTTCTGCTGCTCACCATGCTCCATTATG 1856
Qy 548 LeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCys 567
Db 1857 CTGAATCTGGATTTGTGATCACTACATTAACAATAGCAAGGCAATCAGGAGGATGTGC 1916
Qy 568 PheLeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPhe 587
Db 1917 TTTATGATGATCCTCAGGTGGGAGAAAGTGTGTATGTGCCAATTCCTCCAGAGATT 1976
Qy 588 AspGlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsn 607
Db 1977 GATGGTATTCATCCCAATGACCGTTACGCCAATCGAACACCGTATTCTTTGATATCAAC 2036
Qy 608 MetLysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArg 627
Db 2037 ATGAAAGTCTCGGATGGAATTCAGAGGCTGTATATGTGGGAACCTGGATGATGTTCAGA 2096
Qy 628 ArgGlnAlaLeuTyrGlyTyrAsnProLysGlyProLysArgProLysMetValThr 647
Db 2097 AGACAAGCTCTATATGGGTATGGGCTCCCAAGGCCCAAAACGTCCTCCAGATGGTGACC 2156
Qy 648 CysAspCysCysProCysPheGlyArgLysLysArgLysHisAlaLys----- 663
Db 2157 TGTGATTTGCTCCCTTGTGTGGGTCCTCGTAAGAAGTCTCCGAGAAATAATAGTAGCAAG 2216
Qy 664 -----AspGlyLeuProGluGlyThrAla 671
Db 2217 AAAAGTCGAGGAATCCAGCTCCCGCTACCAATCTGGACGGGATCGAGNAGGATAGAA 2276
Qy 672 AspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArg 691
Db 2277 -----GGTTATGATGACGAAAGAGCATTTGTTGATGAGCCAACTAGACTTCGAGAGAG 2330
Qy 692 PheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGluGlyValProPro 711
Db 2331 TTTGGCCAGTCTTCAGCTTTTGTTCATCCACTCTGTATGGAGAAATGGTGGTGTTCGCGAA 2390
Qy 712 SerSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGlu 731
Db 2391 ACAGCAATCCAGCTGAATTTGTTGAAGAGGCTATTTCATGTATCAGCTGTGGATGATAA 2450
Qy 732 AspLysThrAspTrpGlyLeuGluLeuGlyTyrIleTyrGlySerIleThrGluAspIle 751
Db 2451 GACAAAACGGAAATGGGAAAAGAGCTTGGATGGATCTATGGATCAGTCACAGAGGACATT 2510

QY	752	LeuThrGlyPheLysMetHisCysArgGlyTTPArgSerValTyrCysMetProLysArg	771
DB	2511	CTGACTGGATTCAAGATGCACATCGAGGCTGGCGGTCCATTACTGTATGCCCAACGA	2570
QY	772	AlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAenGlnValLeuArg	791
DB	2571	GCAGCATTCAAAGGTCCTGCTCCNATCAATCTATCAGACCGTTTGAACCAAGGTGTGCGT	2630
QY	792	TrpAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyr	811
DB	2631	TGGGCTTTGGGATCAGTAGAAATTTTTCATGAGCAGACATTGCCCAATCTGGTATGGCTAT	2690
QY	812	LysAsnGlyAsnLeuLysTyrProLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro	831
DB	2691	---GGGGGAGGCTCGAAATGGCTTGAAGAATTTGGCTATATCAACCAATGTGCTATCCA	2747
QY	832	PheThrSerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGly	851
DB	2748	TTCACTCTCTTCCACTCATTTGGCTATTGGACACTTCCAGCCGTCAGTTTGTCTCATCTGGC	2807
QY	852	LysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMet	871
DB	2808	AAATTTGTGATCCTCCAGATCAGTACTTTTGGCAAGTCTATTTTAAATAGCTCTTTTCAATC	2867
QY	872	SerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValSerIleGluGluTyr	891
DB	2868	TCAATTTTTCGCCACTGGTATTTCTGAAATCAGGTGGAGTGGAGTCGAGCATTTGAAGAATGG	2927
QY	892	TrpArgAsnGluGlnPheTyrValIleGlyGlyValSerAlaHisLeuPheAlaVal	911
DB	2928	TGGCGAAATGAACAGTTCTGGGTATTGGAGGGGTTTCTGCACATTTTTTTGCGAGTATT	2987
QY	912	GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla	931
DB	2988	CAAGGTCGTCTCAAGGTACTGGCAGGCATTGATACAAATTTTCAGTCTCATGCGCAAGGCA	3047
QY	932	ThrGlyAspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTyrThrThrLeuLeu	951
DB	3048	TCA-----GATCAGCGTGAGTTTGGGGAACTGTATGCATVTCAAATGGACACACATCCTC	3101
QY	952	IleProProThrThrLeuLeuIleIleAsnValIleGlyValValAlaGlyIleSerAsp	971
DB	3102	ATTCTCTCTACAAACCTTCCTGTGTCATACACTTGTGGGGGTGGTGTGGTGGGTAGCAGAT	3161
QY	972	AlaIleAsnAsnGlyTyrGlnSerTyrGlyProLeuPheGlyLysLeuPhePheAlaPhe	991
DB	3162	GCAATCAACAATCGAATTCAGTCATGCGGTCTCTCTTGGGTAAGCTTTTCTTTGTCATTCT	3221
QY	992	TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr	1011
DB	3222	TGGGTCATTGTGCACCTGTATCTCTTTCTCAAGGGTCTCATGGCGCAGGACGAGAACCA	3281
QY	1012	ProThrValValValIleTyrSerIleLeuLeuAlaSerIlePheSerLeuLeuTyrVal	1031
DB	3282	CCACCATCTGGGTATTGGTCAATTCCTGGCATCTGTTTCTCTCTTTTCTCTGGGTA	3341
QY	1032	ArgIleAspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIleAsn	1051
DB	3342	AGAATTGATCTTCTTGAGTAAGGTTAAGGCCGAGATACTAAACAATGTGCATCAAC	3401
QY	1052	Cys	1052
DB	3402	TGC	3404

RESULT 8

RESULTS 8
US-10-425-114-833

US-10-423-114-833
; Sequence 833, Application US/10425114

;; SEQUENCE 833, APPLICATION US/10
;; PUBLICATION No. US20040034888A1

; FUNDING NO: US20
; GENERAL INFORMATION:
; GENERAL INFORMATION:

APPLICANT: Liu, Jindong

APPLICANT: LIU, Jingdong
APPLICANT: Zhou. Yihua

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic David K

APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.

```

; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 833
; LENGTH: 2687
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700093978_FLI
US-10-425-114-833

```

Alignment Scores:

Pred. No.:	0	Length:	2687
Score:	4395.00	Matches:	824
Percent Similarity:	99.88%	Conservative:	0
Best Local Similarity:	99.88%	Mismatches:	1
Query Match:	77.82%	Indels:	0
DB:	17	Gaps:	0

US-10-627-132-30 (1-1052) X US-10-425-114-833 (1-2687)

228	Qy	GlyGlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArg	247
1	Db	GGCGGGCGCCGATCCCGAAGACATGGACGCCGACGTGGCACGTGAACGACGAGCGAGG	60
248	Qy	GlnProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetVal	267
61	Db	CAGCCGCTGTTCGAGGAAGGTGTGCATCGCGTCGAGCAAGGTGAACCCGTACCGGATGGT	120
268	Qy	IleValValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProVal	287
121	Db	ATCGTGGTGGCTCTCGTTGTGCTCGCCTTCTTCTCCGGTACCCTGATCTGCACCCCGTC	180
288	Qy	ProAspAlaIleGlyLeuTyrLeuValSerIleIleCysGluIleTyrPheAlaIleSer	307
181	Db	CCGGACGCCATCGGGCTGTGGCTCGTCTCCATCATCTCGAGATCTGGTTCGCCATCTCC	240
308	Qy	TyrIleLeuAspGlnPheProLysTyrPheProIleAspArgGluThrTyrIleuAspArg	327
241	Db	TGGATCTCCACCAAGTTCCCAAGTGTTCGCCATCGACCCTCGAGACGTACCTCGACCGC	300
328	Qy	LeuSerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPhe	347
301	Db	CTCTCCCTCAGGTACGAGAGGAAGGGAGCGCGTCTGCTGTCTGGCGGTGACCTGTTC	360
348	Qy	ValSerThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSer	367
361	Db	GTGAGCACGGTGGACCCGCTCAAGGACCGCCGCTGGTGACCGCCAAACACCGTGTCTCC	420
368	Qy	IleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAla	387
421	Db	ATCCTCGCCGTAGACTACCCCGTGGACAAGGTCTCTGCTACGTCTCCGACGACGCGCGC	480
388	Qy	SerMetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValPro	407
481	Db	TCGATGTGACGTTTCGAGTTCGCTGTTCGGAGACGGCCGAGTTTCGGCGCGCAAGTGGGTGCC	540
408	Qy	PheCysValLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysVal	427
541	Db	TTCTGCAAGAAGTTTCGGCATCTGAGCCCGCCGCGCGGAGTTTCTACTTCTCGCTCAAGGTTC	600
428	Qy	AspTyrLeuLysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArg	447
601	Db	GACTACTCTCAAGGACAAGGTTCAGCCACCTTCTGTCAGGAGCGCGCGCCATGAAGAGA	660
448	Qy	GluTyrGluGlnPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValPro	467
661	Db	GAGTATCAGAGATTCAAGTCCGATCAACGCTGTGGTGGCCAGGCGCATGAAGTGGCG	720

QY 468 AlaGluGlyTrpIleMetLysAspGlyThrProTrpProGlyAenAenThrArgAspHis 467
Db 721 GCAGAGGGGGTGAATCATGAAGAGCGGCACCGCTGGCCCGGGAACAACACCCCGGACCAC 780
QY 488 ProGlyMetIleGlnValPheLeuGlyHisSerGlyVlyHisAspThrGluGlyAenGlu 507
Db 781 CCGGGCATGATCCAGGTGTTCTTGGGCGACAGCGGGCGGCCACACACCGAGGGGCACGAG 840
QY 508 LeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLys 527
Db 841 CTGCCCCGCTCGTGTACGTCTCCGTGAGAGCGCCCGGGATTTAGCACCAACCAAG 900
QY 528 AlaGlyAlaMetAenAlaLeuIleArgValSerAlaValLeuAenAlaProPheMet 547
Db 901 GCCGGCGCCATGAACGCTCTGATTCTCGGCTCTCGGCGCTGACCAACCGCCATTTCATG 960
QY 548 LeuAenLeuAspCysAspHisTyrIleAenAenSerLysAlaIleArgGluAlaMetCys 567
Db 961 CTCACTTGGACTGTGATCACTACATCAACACAGCAAGGCCATCCGGAGGCCATGTGC 1020
QY 568 PheLeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPhe 587
Db 1021 TTCTCTATGACCTCAGGTCGGCGGGAAGGTCTGTCTACGTTTCAGTTCCCGCAGAGGTTTC 1080
QY 588 AspGlyIleAspValHisAspArgTyrAlaAenAenAenThrValPhePheAspIleAen 607
Db 1081 GACGGCATCGACGTGACGACCGATACATCAACACAGCAAGGCCATCCCGTCTTTCGACATCAAC 1140
QY 608 MetLysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArg 627
Db 1141 ATGAAGGGGCTGGACGGCATCCAGGCGCCGTGTACGTCTGGGACAGGGTTCGGTTCGGG 1200
QY 628 ArgGlnAlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThr 647
Db 1201 CGCCAGGCGCTCTACGGCTACAACCCCTCCCAAGGACCCCAAGAGGCCCAAGATGGTGACC 1260
QY 648 CysAspCysCysProCysPheGlyVArgLysArgLysHisAlaLysAspGlyLeuPro 667
Db 1261 TGGCACTGCTCCCGTGTCTCGGCGCAAGAGCGGAAACACGCAAGGACGGGCTGCGG 1320
QY 668 GluGlyThrAlaAspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsn 687
Db 1321 GAGGGCACCGCTGATATGGGAGTAGATAGCACCAAGGAGATGCTCATGTCTCCACATGAAC 1380
QY 688 PheGluLysArgPheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGluGly 707
Db 1381 TTCGAGAAGCGGTTCCGGGACGTCCCGCGCGTTCGTACGTCCAGCGCTGATGGAGGAAGGC 1440
QY 708 GlyValProProSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSer 727
Db 1441 GCGCTCCCTCTCTGTCTGAGCGCCCGCGCTCTCAAGAGGGCCATCATGTGTATCAGC 1500
QY 728 CysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeuGlyTyrIleTyrGlySerIle 747
Db 1501 TCGGCTACAGGACACAGACCGACTGGGGGCTGGAGCTGGGGTGATCTACGGGTCTGATC 1560
QY 748 ThrGluAspIleLeuThrGlyPheLysMetHisCysArgGlyTyrArgSerValTyrCys 767
Db 1561 ACGGAGGACATCTCTACCGGGTTCAAGATGCACTGCCCGGGGTGGCGCTCCGTGTACTGC 1620
QY 768 MetProLysArgAlaAlaPheLysGlySerAlaProIleAenLeuSerAspArgLeuAen 787
Db 1621 ATGCCGAAGCGGGCGGCTTCAAGGGGTTCGCGCGCGATCAATCTATCGGACCGTCTCAAC 1680
QY 788 GlnValLeuArgTyrAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeu 807
Db 1681 CAGGTGCTCCGGTGGGCGCTGGGGTTCGCTCGAGATCTTCTTCAGCGGCGACAGCCCGCTG 1740
QY 808 LeuTyrGlyTyrLysAenGlyAenLeuLysTrpLeuGluArgPheAlaTyrIleAenThr 827
Db 1741 CTGTACGGCTACAAGAACGGCAACCTCAAGTGGCTGGAGCGCTTCGCGCTATCAACACC 1800

QY 828 ThrIleTyrProPheThrSerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCys 847
Db 1801 ACCATCTACCCCTTACCTCGCTCCGCTGCTCGCTACTGCAACCTCCCGCGCTGTC 1860
QY 848 LeuLeuThrGlyLysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIle 867
Db 1861 CTCTCTACCGGCAAGTTTCATTCGCGTGCATTTAGCACGTTTCGCCAGCCTCTTCTTCATC 1920
QY 868 AlaLeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValSer 887
Db 1921 GCCCTCTTTCATGTCCATCTTCGCGACGGGATCTCTGGAGATGCGGTGGAGCGGGTGAGC 1980
QY 888 IleGluGluTrpTrpArgAenGluGlnPheTrpValIleGlyValSerAlaHisLeu 907
Db 1981 ATCGAGGAGTGGTGGAGGAACGAGCAGTTCTGGGTTCATCGCGCGCTGTCCGCGCATCTC 2040
QY 908 PheAlaValValGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAenPheThrVal 927
Db 2041 TTCCGCGCTCGTGCAGGGCCTGCTCAAGGTCTCTCGCGGGATCGACCACTTCACCGTC 2100
QY 928 ThrSerLysAlaThrGlyAspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrp 947
Db 2101 ACCTCAAGGCCACCGCGGACGAGCAGGTTCGCCGAGCTCTACGCTTCNAAGTGG 2160
QY 948 ThrThrLeuLeuIleProProThrThrLeuLeuIleIleAenValIleGlyValValAla 967
Db 2161 ACCAGCTCTCTATCCCGCCACCGCTCTCATCATTAACGTCTATCGCGCTCGTGGCC 2220
QY 968 GlyIleSerAspAlaIleAenAenGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeu 987
Db 2221 GGCATCTCCAGCCCATCAACAACGGGTACAGTCTCTGGGGGCGCTCTTCGGCAAGCTC 2280
QY 988 PhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArg 1007
Db 2281 TTCTTCGCTCTCTGGGTTCATCTCCACCTCTACCCGTTCTCAAGGGGCTCATGGGGCGC 2340
QY 1008 GlnAenArgThrProThrValValIleTrpSerIleLeuLeuAlaSerIlePheSer 1027
Db 2341 CAGNACAGGACGCCACCGTTGTTGTCATCTGGTCCATCTGTGTGGCTCTCCATCTTCTCC 2400
QY 1028 LeuLeuTrpValArgIleAspProPheIleValArgThrLysGlyProAspValArgGln 1047
Db 2401 CTGCTCTGGGTTCAGGATCGACCCCTTCATCTCGTAGGACCAAGGGGCGCGAGCTCAGGCAG 2460
QY 1048 CysGlyIleAenCys 1052
Db 2461 TGTGGCATCAATTGC 2475

RESULT 9
US-09-900-237-3
; Sequence 3, Application US/09900237
; Patent No. US200201201241
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant
; FILE REFERENCE: B81170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 3776
; TYPE: DNA
; ORGANISM: Zea mays
US-09-900-237-3
Alignment Scores:

Pred. No.: 0 Length: 3776
 Score: 4091.00 Matches: 760
 Percent Similarity: 79.64% Conservative: 128
 Best Local Similarity: 68.16% Mismatches: 137
 Query Match: 72.43% Indels: 90
 DB: 9 Gaps: 16

US-10-627-132-30 (1-1052) x US-09-900-237-3 (1-3776)

Qy	1	MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGlnLeuValLeu	20
Db	179	ATGAGGCGCAGCGCGCGGTGTCGCGCGCTCGCACACCGGACGAGCTGTCGTCATC	238
Qy	21	Arg-----GlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGlu	38
Db	239	CGCCCGCATGGCGAGCCGAGCGCGGAGCCCATGTGACACGAGCGCGCGGTGTGCCAG	298
Qy	39	IleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGlu	58
Db	299	ATTTCGGGGACAGCTGGGGCGCAACCCGACGGGGAGCCGTTCTGGCCCTGCCAACGAG	358
Qy	59	CysGlyPheProValCysArgProCysTyrGluTyrGluArgGluGluGlyThrGlnAsn	78
Db	359	TGCGCTTCCCATCTGCGGAGCTGCTACGAGTACGAGCGCGCGAGGGCAGCGCAGAAC	418
Qy	79	CysProGlnCysLeuThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAsp	98
Db	419	TGCCCCCAGTGCACACCGCTTCAAGCGCTTCAAGGGGTGCGCGCGGTGCCCGGGAC	478
Qy	99	AspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGln	118
Db	479	GAGGGAGGAGCGCGCTGCACACCTGGAGAACGAGTTCACTGGAGCGACGACGACGAC	538
Qy	119	GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly	138
Db	539	TCCCAGTACCTC-----GCCGAGTCCCATGCTCCACGCC	571
Qy	139	ArgMetSerTyrGlyArgGlyProAsp---AspGlyAspGlyAsnAsnThrProGlnIle	157
Db	572	CACATGAGCTAGCGCGCGCGCCGACCTCGACGGCGTCCGACGCCATTCCACCCCATC	631
Qy	158	ProProlIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsn	177
Db	632	CCC-----AATGTTCCTC-----CTCACCAAC	655
Qy	178	GlyTyrGlyHisGlyGluValSerSerLeuHis-----	189
Db	656	GGACAGATGGTCGATGACATCCGCGCGGACGACGCGCTTGTGCCCTCGTTCGTGGGT	715
Qy	190	-----LysArgIleHisProTyrProValSerGluProGly-----Ser	202
Db	716	GCGCGGGGGAAGAGGATTCACCTCTCCCGTACGCGGATCCCAACCTTCTCTGTGCAACCG	775
Qy	203	AlaLysTrpAspGluLysGlu-----ValSerTrpLys	214
Db	776	AGGCTATGGACCTTCCCAAGATCTCCCGCATATGCTACGAGCGGTAGCTGGAAG	835
Qy	215	GluArgMetAspAspTrpLysSerLysGlnGlyIleLeu-----Gly	228
Db	836	GAGAGGATGGAGAGCTGGAAGCAGAGCAGGAGGAGGATGCACGACGAGAACGATGGC	895
Qy	229	GlyGlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGln	248
Db	896	GCGCGCGATGATGGTGTAT-----GATGCAGATCTACCACCTAAATGGATGAAGCTAGACAG	949
Qy	249	ProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIle	268
Db	950	CCATTGTCCAGAAAGATCCCGCTTCCTTCAAGCCAAATCAACCCCTATAGATGATTATA	1009
Qy	269	ValValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValPro	288
Db	1010	ATAATTGGCTAGTGTGTTGTTGTTCTTCTTCCACTACCGAGTATGATCCATCCGGTGCCT	1069

Qy	289	AspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrp	308
Db	1070	GATGCAATTTGCTTTATGCTCATATCTGTGATCTGTGAAATTTGGTTTGGCCATCTTGG	1129
Qy	309	IleLeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeu	328
Db	1130	ATTCTTGACCAAGTTTCCAAAGTGGTTTCTATCGAGAGGAAACCTATCTTCCACCGGTG	1189
Qy	329	SerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheVal	348
Db	1190	AGTTTAAGGTTTGACAAGAAAGGCGATCCTTCTCAACTCGCCCTGTTGATTTCTTGTTC	1249
Qy	349	SerThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIle	368
Db	1250	AGTAGCGTTGATCCCTTGAAGAAACCTCCATTGGTCACTGCTAACTACTGTTCTATCTATC	1309
Qy	369	LeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSer	388
Db	1310	CTTTGCGTGGATTATCCAGTTGATAAGGTTTATGCTACGTTCTGTGATGGTCTGCC	1369
Qy	389	MetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPhe	408
Db	1370	ATGCTGACATTTGAAGCATTGCTGAAACATCTGAAATTTGCAAGAAATGGTTCCTTTC	1429
Qy	409	CysLysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAsp	428
Db	1430	TGCAAAAGATATAGCTTTGAGCCTCGTCTCCAGAGTGGTACTTCCAAACAGAGATAGAC	1489
Qy	429	TyrLeuLysAspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGlu	448
Db	1490	TACCTGAAAGACAAGGTGGCGCCAACTTTGTTAGAACCGGAGGCAATGAAGAGAGAG	1549
Qy	449	TyrGluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAla	468
Db	1550	TATGAGGAATTCAGGTCAGAAATCAATGCCCTTGGTTGTTAAAGCCCAAGGTTCTCTGAG	1609
Qy	469	GluGlyTrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisPro	488
Db	1610	GAAGGATGGACAATGCAGGATGGAACCTCATGGCCGCGAAATAATGTCCGTGATCATCT	1669
Qy	489	GlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeu	508
Db	1670	GGAATGATTCAGGTTTTCCTTGGTCAAGTGGTGGCCATGATGTGGAAGGAAATGAGCTG	1729
Qy	509	ProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisIleLysLysAla	528
Db	1730	CCTCGATTGTTTATGTTTCAAGAGAAACCGCCAGGCTACACCATCACAGAGGCT	1789
Qy	529	GlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeu	548
Db	1790	GGTCTATGAATGATGGTCCGAGTCTCTGCTGTAATACTAACTATGCTCTTATTTGCTG	1849
Qy	549	AsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPhe	568
Db	1850	AACTTGATTTGATGATCACTATATCAATAATAGTAAGGCTATAAAGGAAGCAATGTGTTT	1909
Qy	569	LeuMetAspProGlnValGlyValGlyLysValCysTyrValGlnPheProGlnArgPheAsp	588
Db	1910	ATGATGATCCTTTGCTTGGAAAGAAAGTTTGTCTATGTGTCAGTTTCTCTCAAGATTGAT	1969
Qy	589	GlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMet	608
Db	1970	GGGATGATGCCATGATCGATATGCTAAACAGAAATGTTGTCTTTTTCGATATCAACATG	2029
Qy	609	LysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArg	628
Db	2030	AAAGTTTGGATGATGATCCAGGCGCCCAATTTATGTGGTACTGTGATGTCTTTCAGAGG	2089
Qy	629	GlnAlaLeuTyrGlyTyrAsnProLysArgProLysArgProLysMetValThrCys	648
Db	2090	CAGGCATATATGCTACGATGCTCCCAAA---ACAAGAGAGCCACCATCAAGAACTTGC	2146
Qy	649	Asp-----CysCysProCysPheGly---ArgLysLysArgLys	660

QY 119 GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
Db 444 TCCAGTACCTC-----GCCAGTCCATGCTCCACGCC 476
QY 139 ArgMetSerTyrGlyArgGlyProAsp---AspGlyAspGlyAsnAsnThrProGlnIle 157
Db 477 CACATGAGTACGCCCGCGCGCGACCTCGACGGGTGCCGAGCCATCCACCCATC 536
QY 158 ProProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsn 177
Db 537 CCC-----AATGTTCCCTC-----CTCACCAAC 560
QY 178 GlyTyrGlyHisGlyGluValSerSerSerLeuHis----- 189
Db 561 GGACAGATGGTCATGACATCCCGCGGACACGACGCCCTGTGTGCCCTCGTTCGGGT 620
QY 190 -----LysArgIleHisProTyrProValSerGluProGly-----Ser 202
Db 621 GCGCGGGGAGAGGATTCACCTCTCCCGTAGCGGGATCCCAACCTTCTGTGCACCG 680
QY 203 AlaLysTrpAspGluLysLysGlu-----ValSerTrpLys 214
Db 681 AGGTCTATGGACCTTCCCAAGGATCTCGCGCATATGGCTACGGGAGCGTAGCATGGAAG 740
QY 215 GluArgMetAspAspTrpLysSerLysGlnGlyIleLeu-----Gly 228
Db 741 GAGAGGATGGAGCTGGAGGAGCAGAGCAGAGGAGATCGACACGACGAGAACGATGCG 800
QY 229 GlyGlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGln 248
Db 801 GCGCGGATGATGGTAT-----GATGCAGATCTACCACTAATGGATGAAGTACAG 854
QY 249 ProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIle 268
Db 855 CCATTTGTCAGAAAGATCCGCTTCCTTCAAGCCAAATCAACCCCTATAGATGATTATA 914
QY 269 ValValArgLeuValValLeuAlaPheLeuArgTyrArgIleLeuHisProValPro 288
Db 915 APTAATTCGGCTAGTGGTTGTGTTCTTCTTCCACTACCGAGTGTGATCGCGTGCCT 974
QY 289 AspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrp 308
Db 975 GATGATTTGCTTTATGGCTATATCTGTGATCTGTGAATTTGGTTGCCATGCTTGG 1034
QY 309 IleLeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeu 328
Db 1035 ATTTCTGACAGTTTCCAAAGTGTTCCTATCGAGAGGAAACCTAATCTTGACCGCTG 1094
QY 329 SerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheVal 348
Db 1095 AGTTTAAAGGTTTGACAAAGGAGGGGATCCTTCTCAACTCGCCCTGTGATTTCTTGTG 1154
QY 349 SerThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIle 368
Db 1155 AGTACGGTTGATCCCTTGAGAGAACTCATTTGGTCACTGCTAATACTGTTCTATCTATC 1214
QY 369 LeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAlaSer 388
Db 1215 CTTTCGGTGGATTATCCAGTTGATTAAGTTTCATGCTACGTTCTTGATGATGGTCTGCC 1274
QY 389 MetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPhe 408
Db 1275 ATGCTGACATTTGAAGCATTTGTGAAACATCTGAATTTGCAAGAAATGGGTTCCCTTC 1334
QY 409 CysLysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAsp 428
Db 1335 TGCAAAAGATATAGCCTTGAGCCTCGTCTCCAGAGTGGTACTTCCAAACAGAGATGAC 1394
QY 429 TyrLeuLysAspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGlu 448
Db 1395 TACCTGAAAGACAGGTGCGCCAAACTTTTGTAGAGAACGAGGAGCAATGAGAGAG 1454
QY 449 TyrGluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAla 468

Db 1455 TATGAGGAATTCAGGTCAAGATCAATGCTTGTGTTGCTTAAAGCCCCAAAAGGTTCTCTGAG 1514
QY 469 GluGlyTrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisPro 488
Db 1515 GAAGATGACAACTGAGTGGAACTCCATGCCCGGAAATAAATTCGCGTATCATCTT 1574
QY 489 GlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeu 508
Db 1575 GGAATGATTCAGGTTTCTGCTGCTCAAGTGTGGCCATGATGTGAAGAAATGAGCTG 1634
QY 509 ProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAla 528
Db 1635 CCTCATTTGGTTTATGTTTCAAGAGAAAACGCCGAGGCTACACCATCACAGAGGCT 1694
QY 529 GlyAlaMetAsnAlaLeuLeuArgValSerAlaValLeuThrAsnAlaProPheMetLeu 548
Db 1695 GGTGCTATGAATGATTCATGTCGAGTCTCTGCTGTAACATAATGCTCTCTTATTGCTG 1754
QY 549 AsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPhe 568
Db 1755 AACTTTGGATTTGATCACATATATCAATAGTAGGCTATAAAGAGCAATGTGTTTT 1814
QY 569 LeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAsp 588
Db 1815 ATGATGATCCCTTCTTGGAAATAAGTTCATGTCAGTTCCTCAAGATTTGAT 1874
QY 589 GlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPheAspIleAsnMet 608
Db 1875 GGGATTTGATCGCCATGATCGATATGCTAACAGAAATGTTGCTTTTTCGATATCAACATG 1934
QY 609 LysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArg 628
Db 1935 AAAGTTTGGATGGTATCCAGGGCCCAATTTATGTTGGTACTGGATGTCTTTCAGAGG 1994
QY 629 GlnAlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCys 648
Db 1995 CAGGCATTTATGCTAGCTCTCCCAA---ACAAAGAGACCCACCATCAAGAACTTGC 2051
QY 649 Asp-----CysCysProCysPheGly---ArgLysLysArgLys 660
Db 2052 AACTGCTGGCCAAAGTGGTGCATTTGCTGTTGCTGTTTGGTAAACAGAGACCAAGAG 2111
QY 661 HisAlaLysAspGlyLeuProGlu----- 668
Db 2112 AAGACCAAGACCTCTAAACCTTAATTTGAGAAGATAAAGAACTTTTAAAGAAAAGGNA 2171
QY 669 -----GlyThrAlaAspMetGlyValAspSer 677
Db 2172 AATCAAGCCCTGCATATGCTTCTTGGTGAATTTGATGAAGCTGCTCCAGAGCTGAAAAT 2231
QY 678 AspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAla 697
Db 2232 GAAAGGCTAGTATTGTAAATCAACAGAAAGTTGGAAGAAATTTGGCCAGCTTTCAGTT 2291
QY 698 PheValThrSerThrLeuMetGluGlyGlyValProProSerSerSerProAlaAla 717
Db 2292 TTTGTTGCATCCACATCTTGGAGAAAGTGGTGAACCTTGAAGAGTGCAGCTCCAGCTTCT 2351
QY 718 LeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGly 737
Db 2352 CTTCTGAAGGAAGCTATACATGTCATGCTGTGGATATGAAGACAAACAGGCTGGGA 2411
QY 738 LeuGluLeuGlyTyrIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMet 757
Db 2412 AAAGATATTGGTTGGATTTATGGAATCAGTCACAGAAAGATATTTCTTACGTTTAAAGATG 2471
QY 758 HisCysArgGlyTyrArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySer 777
Db 2472 CACTGCCATGTTGGCGGTCAATTTACTGATACCTAAACGGCGCCCTTCAAGGTTCC 2531
QY 778 AlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerVal 797

Db 2532 GCACCTCTCAATCTTCCGATCGTCTTCCACAGGTTCTTCGGTGGGCTCTTGGTTCAAATT 2591
Qy 798 GluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrIysAsnGlyAsnLeuLys 817
Db 2592 GAAATTTTCTTCAGCAACCACTGCCCTCTCTGGTATGGGTAT---GGTGTGAGCTAAAG 2648
Qy 818 TrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeu 837
Db 2649 TTCTCTGAAAGGTTTTCGTACATTAATCCATCGTATACCTTGGACATCTATCCCGCTC 2708
Qy 838 LeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyIysPheIleMetProSer 857
Db 2709 TTGGCTTATTGCATTTGCCCTGCCATCTGCTGTGACAGGAAATTTATCAGCCACAG 2768
Qy 858 IleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGly 877
Db 2769 CTTAAACAATGTGCGACGCTGTGGTTTCACTTTCATCTGTCATTTTGTGTACGAGC 2828
Qy 878 IleLeuGluMetArgTyrSerGlyValSerIleGluTyrTrpArgAsnGluGlnPhe 897
Db 2829 ATCTTGAAATGAGATGGAGTGTGTAGGCATCGATGCTGTGGAGAAACGACGACTT 2888
Qy 898 TrpValIleGlyIysValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysVal 917
Db 2889 TGGGTCAATTGGAGCGTGTCTTCATCTCTTTGCTGTGTTCCAGGACTCTCTCAAGGTC 2948
Qy 918 LeuAlaGlyIleAspThrAsnPheThrValThrSerIysAlaThrGlyAspGluAspAsp 937
Db 2949 ATAGCTGGTGTAGACACGAGCTTCACCTGTGATCCCAAG-----GGCGGAGACGACGAG 3002
Qy 938 GluPheAlaGluLeuTyrAlaPheIysTrpThrThrLeuLeuIleProProThrThrLeu 957
Db 3003 GAGTTCTCAGAGCTGTATACATTCATATGGAGACCTCTTGATACCTCCGACAAACCTGT 3062
Qy 958 LeuIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyr 977
Db 3063 CTCCTACTGAACCTTCAATTGGAGTGTGTAGCTGGCATCTCCAAATGCGATCAACAACGGATAT 3122
Qy 978 GlnSerTrpGlyProLeuPheGlyIysLeuPhePheAlaPheTrpValIleValHisLeu 997
Db 3123 GAATCATGGGGCCCCCTGTTCGGGAAGCTCTTCTTGTGATTTTGGGTGATCGTCCATCTT 3182
Qy 998 TyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValValIle 1017
Db 3183 TACCGTTCTCTCAAGGTCGTGGTGGAGGAGGAGACAGACGCCAACGATTGTATTGTC 3242
Qy 1018 TrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIle 1037
Db 3243 TGGTCCATCTCTCGCTTCGATCTTCTGCTGCTTTGGGTCCGATCGACCCGTTCTT 3302
Qy 1038 ValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3303 GCGAAGGATGATGGTCCCTCTGTGGAGGATGTGTGTGGATTGC 3347

RESULT 11

US-10-425-115-86998
; Sequence 86998, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 86998
; LENGTH: 3783
; TYPE: DNA
; ORGANISM: Zea mays

FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_179348C.1

US-10-425-115-86998

Alignment Scores:
Pred. No.: 0 Length: 3783
Score: 4089.00 Matches: 760
Percent Similarity: 79.55% Conservative: 127
Best Local Similarity: 68.16% Mismatches: 138
Query Match: 72.40% Indels: 90
DB: 18 Gaps: 16

US-10-627-132-30 (1-1052) x US-10-425-115-86998 (1-3783)

Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
Db 186 ATGAGGCGCAGCCCGGGCTGTGCGGGCTCGCACACACCGAAGACGAGCTCGTCGTCATC 245
Qy 21 Arg-----GlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGlu 38
Db 246 CGCCGCGATGGCGAGCCAGGGCCGAAGCCCATGCAGCCGGAACGCGCAGGTGTGCCAG 305
Qy 39 IleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGlu 58
Db 306 ATTTGGCGCAGCAGCTGGGGCGCAACCCCGACGGGAGCCGTTCTGTGGCTGTGCAACGAG 365
Qy 59 CysGlyPheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsn 78
Db 366 TGGCCTTCCCAATCTGCGGGACTGCTACAGTACGAGCGCCGCGAGGCGCAGCAGAAC 425
Qy 79 CysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAsp 98
Db 426 TGCCCCCAGTGCAAGACCCGCTTCAAGGGGTGCGCGCGCTGCCCGGGAC 485
Qy 99 AspAspGluIleAspLeuAspLeuGluHisGluPheAsnIleAspAspGluAsnGln 118
Db 486 GAGGAGGAGGAGCGCGCTCGACACCTCGAGAGGTTCACTGGACGCAACGACGAC 545
Qy 119 GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
Db 546 TCCAGTACCTC-----GCCGAGTCCATGCTCCACGCC 578
Qy 139 ArgMetSerTyrGlyArgGlyProAsp---AspGlyAspGlyAsnAsnThrProGlnIle 157
Db 579 CACATGAGTTACGCGCGCGCGACCTCGACGCGCTGCGCGACCATTCACCCCATC 638
Qy 158 ProProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsn 177
Db 639 CCC-----AATGTTCCCTC-----CTCACCAAC 662
Qy 178 GlyTyrGlyHisGlyGluValSerSerSerLeuHis----- 189
Db 663 GGACAGATGTCGATGACATCCCGCCGAGCAGCAGCCCTTGTGCTCGTGGGT 722
Qy 190 -----LysArgIleHisProTyrProValSerGluProGly-----Ser 202
Db 723 GGCGGGGGAAGAGATTACCTCTCCGTCACGCGATCCCACTTCCTCTGTGCAACCG 782
Qy 203 AlaLysTrpAspGluLysLysGlu-----ValSerTrpLys 214
Db 783 AGGTCTATGGACCTTCCAAAGGATCTCGCCGCATATGGCTACGGAGCGTAGCATGGAAG 842
Qy 215 GluArgMetAspAspTrpLysSerLysGlnGlyIleLeu-----Gly 228
Db 843 GAGAGATGGAGAGCTGGACGACAGCAGGAGGAGGATGACACGAGGAGGACGATGCG 902
Qy 229 GlyGlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGln 248
Db 903 GGCGCGCATGATGTGAT-----GATGCAGATCTACCACTTAATGATGAAGTAGACAG 956
Qy 249 ProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIle 268
Db 957 CCATTGTCCAGAAAGATCCCGCTTCTTCAAGCCAAATCAACCCCTTAGATGATTATA 1016

Db 1455 TTGAAGACAGGTGGCAGCAAACTTTGTTAGGAGAGAGAGCAATGAAGAGAGATAT 1514
Qy 450 GluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGlu 469
Db 1515 GAGGAATTCAGAGTGAAGATCAATGCTTAGTTGCCAAGAGCCAGAGAGTCTCTAGAA 1574
Qy 470 GlyTrrPheMetLysAspGlyThrProTrrProGlyAsnAsnThrArgAspHisProGly 489
Db 1575 GGATGGACAATCAAGATGGAAACCCCTGGCCCTGGAAACAATGTTGCTGATCATCTGGA 1634
Qy 490 MetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGluAsnGluLeuPro 509
Db 1635 ATGATTGAGTCTCTCTGGCCAAAGCGGAGCCCTTGACTGTGAGGAGAAATGAATGGCCA 1694
Qy 510 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisIleLysLysAlaGly 529
Db 1695 CGATTGGTTTATGTTCTTAGAGAGAAACGACCGAGCTATAACCATCAATGAAGAAGCTGGT 1754
Qy 530 AlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsn 549
Db 1755 GCTATGAATGCATGGCTCGAGTCTCTGCTGTAACAACAATGCTCCATATTTGTTAAAC 1814
Qy 550 LeuAspCysAspHisTrrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeu 569
Db 1815 TTGGATTGTGATCACTACATCAACACAGCAAGGCTATAAAGGAAGCAATGTTTATG 1874
Qy 570 MetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGly 589
Db 1875 ATGGACCCCTTTACTAGGAAGAAGGCTTGTCTATGTACAGTTCCTCAAGATTTGATGG 1934
Qy 590 IleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLys 609
Db 1935 ATTGATCCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTCATATCAACATGAA 1994
Qy 610 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgGln 629
Db 1995 GATTGGATGGTATTCAGGCTCCAATTTATGTTGTACTGGATGTGTTATAGAAAGGCAG 2054
Qy 630 AlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysAsp 649
Db 2055 GCATTATATGTTATGATGCCCCCAA--ACAAAGAGCCACCATCAAGGCTTGCAAC 2111
Qy 650 -----CysCysProCysPheGlyArgLysLysArgLysHisAla 662
Db 2112 TGCTGCCCAAGTGTGCTTTTGTCTGCTTGTGCTTGGCAATAGGAAGCAAGAAAGACT 2171
Qy 663 LysAspGlyLeuProGlu----- 668
Db 2172 ACCAAACCAACAGAGAGAAAGTATATTTTTCAGAAAGAGAGAAACCAATCC 2231
Qy 669 -----GlyThrAlaAspMetGlyValAspSerAspLysGlu 680
Db 2232 CCTGCATATGCTCTGTGTAATTCACAGAGCTGCTCCAGGAGCTGAGATGAAGAGGCC 2291
Qy 681 MetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThr 700
Db 2292 GGTATTGTAATCAACAAAAATTAGAAAAAATTTGGCCAATCTCTGTTTGTGTTACA 2351
Qy 701 SerThrLeuMetGluGlyGlyValProProSerSerSerProAlaAlaLeuLys 720
Db 2352 TCCACACTTCGAGAAATGGTGGAAACCTTGAGAGAGTGAAGTCTGCTCTCTTTTGAA 2411
Qy 721 GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrrPheGluGluLeu 740
Db 2412 GAAGCTATACATGATCATTTAGTTGTTATGAGACAGACAGACTGGGGAAGAGAT 2471
Qy 741 GlyTrrPheLysSerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg 760
Db 2472 GGCTGGATCTATGATCATGATACAGAGATATCTTAAGTGGTTTCAAGATGCAATGTCAT 2531
Qy 761 GlyTrrArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIle 780

Db 2532 GGTGGCGGTCAATTTACTGTCATACCTAAACGGGTGTCATCAAAAGTTCTCACACTCTG 2591
Qy 781 AsnLeuSerAspArgLeuAsnGlnValLeuArgTrrPheAlaLeuGlySerValGluIlePhe 800
Db 2592 AATCTTTTCAGATCGTCTTCCAGAGTGTCTCGGTGGCTCTTGGGTCTATTGAGATCTTC 2651
Qy 801 PheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrrPheGlu 820
Db 2652 TTCAGCAATCATTCGCTCTTGGTATGGGTAT--GGTGGCGGTCTCGAAATTTTGGAA 2708
Qy 821 ArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyr 840
Db 2709 AGATTTTCTCATCACTCACTCATCTGTCATCTCTGGACATATTTCCCTCTTGGCTTAC 2768
Qy 841 CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr 860
Db 2769 TGTACATTCCTGCCATCTGTTTATTGACAGGGAATTTATCACTCCAGAGCTGAATAAT 2828
Qy 861 PheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGlu 880
Db 2829 GTTGGCAGCTGTGGTTCATGTCATCTTTTATCTGCAATTTTGTCTACGAGCATCTAGAA 2888
Qy 881 MetArgTrrSerGlyValSerIleGluGluTrrPheArgGlnGluPheTrrValIle 900
Db 2889 ATGAGATGGAGTGGTGGTGGAAATGATGACTGGTGGAGGAATGAGAGTCTTGGGTCAAT 2948
Qy 901 GlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAlaGly 920
Db 2949 GGAGGTGTCTCTCACACTCTTGTCTGTCTCCAGGAGACTTCTCAAGGTCTAGCTGGT 3008
Qy 921 IleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspAspGluPheAla 940
Db 3009 GTTGATACAAGCTTCACCGTGACATCAAG-----GGTGGAGATGATGAGGAGTCTCA 3062
Qy 941 GluLeuTyrAlaPheLysTrrThrThrLeuLeuIleProProThrThrLeuLeuIleIle 960
Db 3063 GAGCTATATACATTCAAATGGACTACCTTATTGATACCTCTTACCACCTTGTCTTATTTG 3122
Qy 961 AsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTrr 980
Db 3123 AACTTCATTGTGTGTGGTGGCTGGCTTCAATGGATCAATACGGATATGAGTCATGG 3182
Qy 981 GlyProLeuPheGlyLysLeuPhePheAlaPheTrrValIleValHisLeuTyrProPhe 1000
Db 3183 GGCCCTCTTCTTGGGAAGTATCTTTCATTTTGGTGGTATGTCATCTTATCCCTTT 3242
Qy 1001 LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValValIleTrrSerIle 1020
Db 3243 CTCAAAGTGTGGTGGAAAGCAAAACAGGACCAACGATGTGTCATCGTCTGGTCCAT 3302
Qy 1021 LeuLeuAlaSerIlePheSerLeuLeuTrrValArgIleAspProPheIleValArgThr 1040
Db 3303 CTGCTGGCTTCAATCTCTCGCTCTTGGTTCGGATGATCTTCTTCTTCCGAGGAT 3362
Qy 1041 LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3363 GATGCTCGCTTCTTGGAGAGTGTGGTGGTGGATGTC 3398

RESULT 13

US-10-209-059-49
; Sequence 49, Application US/10209059
; Publication No. US2003016383A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Theoreof
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383

```
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 3969
; TYPE: DNA
; ORGANISM: Zea mays
US-10-209-059-49

Alignment Scores:
Pred. No.: 0 Length: 3969
Score: 4073.50 Matches: 757
Percent Similarity: 79.88% Conservative: 129
Best Local Similarity: 68.08% Mismatches: 139
Query Match: 72.12% Indels: 87
DB: 16 Gaps: 15

US-10-627-132-30 (1-1052) x US-10-209-059-49 (1-3969)
Qy 1 MetGluLaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuLeu 20
Db 144 ATGAGGCGAGCGCGCGCTGGTGGCGGCTCCACCAACCGCAGAGCTCGTCTCATC 203
Qy 21 Arg-----GlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGlu 38
Db 204 CGCCGCGAGCGCGATCCCGGCCGAGCCGCCGCGGAGCAGACGGCGCGATGTGCCAG 263
Qy 39 IleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGlu 58
Db 264 ATTTGGCGGAGCGAGCTGGCTTGCCTCCCGCGGGGAGCCCTTCTGGTGGCGTCAACGAG 323
Qy 59 CysGlyPheProValCysArgProCysTyrgluTyrgluArgArgGluGlyThrGlnAsn 78
Db 324 TCGCGCTTCCCGCTGTCGCGGACTGCTACGAATACGAGCGCGGAGGCGCAGCAAC 383
Qy 79 CysProGlnCysHisThrArgTyrglyArgLeuGlySerProArgValAlaGlyAsp 98
Db 384 TGCCCCAGTGCAAGACTCGATACAGCGCTCAAGGGCTGCCAACGTTGACCGGTGAC 443
Qy 99 AspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGln 118
Db 444 GAGGAGGAGGCGCGCTGATGACTGGCAACAGAGTTCACTGGGAGCGGCCATGACTCG 503
Qy 119 GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
Db 504 CAG-----TCTGTGGCGGAGTCCATGCTCTACGCG 533
Qy 139 ArgMetSerTyrglyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIlePro 158
Db 534 CACATGAGTACGCGCGGTGGA-----GGTGACCTCAATGGCGCGCCACAGCTTTC 584
Qy 159 ProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGly 178
Db 585 CAGCTC-----AACCCCAATGTTCCACTC-----CTCACCAACGGG 620
Qy 179 TyrGlyHisGlyGluValSerSerSerLeuHis----- 189
Db 621 CAAATGGTGGATGACATCCCGGAGCAGCAGCGCGTGGTGCCTTCTTTCATGGGTGGT 680
Qy 190 -----LysArgIleHisProTyrglyProValSerGluProGly-----SerAla 203
Db 681 GGGGAAAGAGAGATACATCCCTTCTTATGGGATCCCGAGCTTACCTGTGTGCAACCCAGG 740
Qy 204 LysTrpAspGluLysGlu-----ValSerTrpLysGlu 215
Db 741 TCTATGGACCCATCCAGAGATCTTCTGCATATGGGTATGGTGTGCTTGGAGGAA 800
Qy 216 ArgMetAspAspTrpLysSerLysGln-----GlyIleLeuGlyGly 229
Db 801 CGGATGGAGAAATTGGAAAGCAGACAGAGAGATGCCACAGCGGGGAATGATGGTGT 860
```

```
Qy 230 GlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnPro 249
Db 861 GGTGATGATGGTGAC-----GATGCTGATCTACCACTAATGATGAAGCAAGCAACA 914
Qy 250 LeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrglyArgMetValIleVal 269
Db 915 CTGTCAGGAAATTCACCTTCATCAAGCCAGATTAATCCATATAGGATGATTATCAT 974
Qy 270 ValArgLeuValValLeuAlaPhePheLeuArgTyrglyArgIleLeuHisProValProAsp 289
Db 975 ATTCGGCTGTGGTGTGGGGTCTTCTTCCATCCAGATGATGATCCGGTGAATGAT 1034
Qy 290 AlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIle 309
Db 1035 GCATTTGCTTTGGTGCATATCTGTATCTGTGAAATCTGGTTTGCATGCTTGGATT 1094
Qy 310 LeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrglyLeuAspArgLeuSer 329
Db 1095 CTTGATCAATTCGCAAGTGGTTCCTTATTTAGAGAGAGACTTACCTAGACCGGCTGTCA 1154
Qy 330 LeuArgTyrgluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSer 349
Db 1155 CTGAGGTTCCAGCAAGGAGGCCCATCTCAACTGCTCCAATGATTTCTTTGTGAGT 1214
Qy 350 ThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeu 369
Db 1215 ACGTTGATCCCTTAAAGGAACCTCCTTGTGTGCACAAATACTGTTCTATCTATCCTT 1274
Qy 370 AlaValAspTyrglyProValAspLysValSerCysTyrglyValSerAspGlyAlaSerMet 389
Db 1275 TCGTGGATATCTCTGTGATAGGTTCTTGTCTATGTTCTGTATGATGGTGTGCAATG 1334
Qy 390 LeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCys 409
Db 1335 CTAACGTTTGAAGCATTTCTGAAACATCTGAAATTTGCAAGAAATGGTTCCTTTCTGC 1394
Qy 410 LysLysPheGlyIleGluProArgAlaProGluPheTyrglyPheSerLeuLysValAspTyrgly 429
Db 1395 AACGGTACATATTGAACCTCGCGCTCCAGAGTGGTACTTCCCAACAGAGATAGACTAC 1454
Qy 430 LeuLysAspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGluTyrgly 449
Db 1455 TTGAAACAGCAAGTGGCAGCAAACTTTGTTAGGAGAGAGAGCAANTGAAGAGAGAT 1514
Qy 450 GluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGlu 469
Db 1515 GAGGAATTCAGGTGAGATCAATGCTTGTAGTCCAAAGCCCAAGAAAGTTCTCTGAAGAA 1574
Qy 470 GlyTrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly 489
Db 1575 GGATGGACATGCAAGATGGAAACCCCTGGCCTGGAAACAATGTTCTGTGATCATCTCTGA 1634
Qy 490 MetIleGlnValPheLeuGlyHisSerGlyHisAspThrGluGlyAsnGluLeuPro 509
Db 1635 ATGATTCAGGTCTCTCTTGGCCAAAGCGGCGCTTGTGAGGGAATAATGAATCTGCA 1694
Qy 510 ArgLeuValTyrglyValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGly 529
Db 1695 CGATTTGGTTTATGTTCTTAGAGAGAAACGACGAGCTATAACCATCATTAAGAAAGCTGT 1754
Qy 530 AlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsn 549
Db 1755 GCTATGAATGATTTGGTCCGAGTCTCTGCTGTACTTAACAAATGCTCCATATTTGTTAAAC 1814
Qy 550 LeuAspCysAspHisTyrglyIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeu 569
Db 1815 TTGGATTGTGATCACTACATCAACAGCAAGGCTATAAAGGAAGCAATGTTGTTTATG 1874
Qy 570 MetAspProGlnValGlyArgLysValCysTyrglyValGlnPheProGlnArgPheAspGly 589
Db 1875 ATGAGCCCTTACTTAGGAAAGAGGTTTGTCTATGTACAGTTCCTCCCAAGATTTGATGGG 1934
Qy 590 IleAspValHisAspArgTyrglyAlaAsnArgAsnThrValPhePheAspIleAsnMetLys 609
```


QY 39 IleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGlu 58
DB 264 ATTTGCCGCGCAGACGTCGGCTTGGCCCGCGGGGACCCCTTCGTGGGTGCAACGAG 323
QY 59 CysGlyPheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsn 78
DB 324 TGGCCTTCCCGCTCGCGGACTGCTACGAATACGAGCGCGGGAGGCGCAGCAAC 383
QY 79 CysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAsp 98
DB 384 TGCCTCCAGTGCAGACTCGATACAGCGCTCAAGGGCTGCCAACGCTGACCGGTGAC 443
QY 99 AspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGln 118
DB 444 GAGGAGGAGGACGGCTGATGACTGCGACMACGAGTTCACTGGGACGGCCATGACTCG 503
QY 119 GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
DB 504 CAG-----TCTGTGGCGGAGTCCATGCTCTACGGC 533
QY 139 ArgMetSerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIlePro 158
DB 534 CACATGACTACGGCGCTGGA-----GGTGACCTTAATGGCGCGCCACAAGCTTTC 584
QY 159 ProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGly 178
DB 585 CAGTCT-----AACCCCAATGTTCCATCT-----CTCACCAACGGG 620
QY 179 TyrGlyHisGlyGluValSerSerLeuHis----- 189
DB 621 CAATGGTGGATGACATCCACCGGAGCAGCAGCGCTGCTCTTTTCATGGGTGGT 680
QY 190 -----LysArgIleHisProTyrProValSerGluProGly-----SerAla 203
DB 681 GGGGAAAGAGAGATACATCCCTTCTTATGCGGATCCCACTTACCTGTGCAACCCAGG 740
QY 204 LysTrpAspGluLysLysGlu-----ValSerTrpLysGlu 215
DB 741 TCTATGGACCCATCCAGGATCTTGCTGCATATGGGTATGTGTGTGGTGGAGGA 800
QY 216 ArgMetAspAspTrpLysSerLysGln-----GlyIleLeuGlyGly 229
DB 801 CGGATGGAGAATTGGAGCAGACAGACAGAGGATGCACAGCGGGGAATGATGGTGGT 860
QY 230 GlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnPro 249
DB 861 GGTGATGATGGTGAC-----GATGCTGATCTACCACCTAATGGATGAAGCAAGCAACA 914
QY 250 LeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleVal 269
DB 915 CTGTCAGGAANAATCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATATCAT 974
QY 270 ValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAsp 289
DB 975 ATTCGGCTTGTTTGGGTTCCTTCCACTACCGAGTGATGATCATCCGGTGAATGAT 1034
QY 290 AlalleGlyLeuTrpLeuValSerIlelleCysGluIleTrpPheAlaIleSerTrpIle 309
DB 1035 GCATTTGCTTTGCTGCATATCTGTATCTGTGAATCTGGTATCTGGTATCTTTGGATT 1094
QY 310 LeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSer 329
DB 1095 CTGTGATCAATCCCAAGAGGTTCCTATTTGAGAGAGACTTACTAGACCGGTGICA 1154
QY 330 LeuArgTyrGluArgGlyGluProSerLeuLeuSerAlaValAspLeuPheValSer 349
DB 1155 CTGAGGTTGCAAGAGAGGCGCCAGCATCTCAACTTGTCTCCCAATGATTTCTTTGTCAGT 1214
QY 350 ThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeu 369
DB 1215 ACGGTGATCCCTTAAAGGAACCTCCTTTGGTCACAAACAATACTGTCTATCTATCCTT 1274
QY 370 AlaValAspTyrProValAspLysValSerCysTyrValSerAspGlyAlaSerMet 389

DB 1275 TCGGTGGATTATCCTGTGATAAGTTTCTTGCTATGTTTCTGATGATGCTGCTCAATG 1334
QY 390 LeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCys 409
DB 1335 CTAACTGTTGAACGATATCTGAAACATCTGAATTTGCAAGAAATGGGTCTCTTTCTGC 1394
QY 410 LysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyr 429
DB 1395 AAACGGTACATATTTGAACTCTCGCTCCAGAGTGGTACTTCCACACAGAAGATAGACTAC 1454
QY 430 LeuLysAspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGluTyr 449
DB 1455 TTGAAGACAAAGTGGCAGCAAACTTTGTTAGGAGAGAGAGCAATGAAGAGAGATAT 1514
QY 450 GluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGlu 469
DB 1515 GAGGAATTCAGGTGAGATCAATGCTTAGTTAGTCCAAAGCCCAAGAAAGTTCTCTGAAGAA 1574
QY 470 GlyTrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly 489
DB 1575 GGATGGCAATGCAGATGGAAACCCCTCGCTGGAAACAAATGTTCTGATCATCTCTGA 1634
QY 490 MetIleGlnValPheLeuGlyHisSerGlyHisAspThrGluGlyAsnGluLeuPro 509
DB 1635 ATGATTGAGTCTTCTTGGCCAAAGCGAGGCTTCACTGCTGAGGGAATGAAGTGGCA 1694
QY 510 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysLysLysLys 529
DB 1695 CGATTGGTTTATGTTCTAGAGAGAAACGACCGGCTATTAACCATCATTAAGAAAGCTGGT 1754
QY 530 AlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsn 549
DB 1755 GCTATGATGATGGTGGTCCGAGTCTCTGCTACTACTACAAATGCTCCATATTTGTTAAAC 1814
QY 550 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeu 569
DB 1815 TTGGATTGTGATCACTACATCAACACAGCAAGGCTATAAAGGAAGCAATGTGTTTATG 1874
QY 570 MetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGly 589
DB 1875 ATGACCCCTTCTACTAGAAAGAAAGGTTTGTCTATGATGATGATGATGATGATGATG 1934
QY 590 IleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLys 609
DB 1935 ATTGATCGCCATGACCGATGCTAAACCGGATGTTGCTCTTTTGTATATCAACATGAAA 1994
QY 610 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGln 629
DB 1995 GGTTCGATGATTAATCAGGCTCAATTTATGTTGTTACTGATGATGATGATGATGATGATG 2054
QY 630 AlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysAsp 649
DB 2055 GCATTTATATGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2111
QY 650 -----CysCysProCysPheGlyArgLysLysLysLysLysLysLysLysLysLys 662
DB 2112 TGCTGGCCCAAGTGGTGTGTTTGTGTTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 2171
QY 663 LysAspGlyLeuProGlu----- 668
DB 2172 ACCAAACCAAAACAGAGAGAAAGTTATTTATTTTCAAGAAAGAGAGAACCAATCC 2231
QY 669 -----GlyThrAlaAspMetGlyValAspSerAspLysGlu 680
DB 2232 CCTGCATATGCTTGTGGTGAATTCAGCAAGTGTCTCCAGGAGCTGAGATGAAGAGGCC 2291
QY 681 MetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThr 700
DB 2292 GGTATTTAAATCAACAAATAATAGAAAGAAATTTGGCCAATCTCTGTTTTTGTATCA 2351
QY 701 SerThrLeuMetGluGlyGlyValProProSerSerSerProAlaLeuLeuLys 720

Db 2352 TCCACACTTCTCGAGAAATGGTGGACCTTGAAGAGTGGACCTCTCTCTTTGAAA 2411
Qy 721 GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeu 740
Db 2412 GAAGCTTATACATGTCATTAAGTGTGTGTTATGAAGACAGACAGACTGGGAAAGAGATT 2471
Qy 741 GlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg 760
Db 2472 GGCTGATCTATGGATCACTTACAGAAAGATATCTAACTGGTTCAAGATGATGTCAT 2531
Qy 761 GlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIle 780
Db 2532 GTTGGCGGTCAATTACTTGCATACCTAAACGGGTGTCATTCAAGGTTCTGCACCTCG 2591
Qy 781 AsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePhe 800
Db 2592 AATCTTTTCAGATCGCTTCCACAGGTGCTTCGGTGGCTCTTGGGTCTATTGAGATCTTC 2651
Qy 801 PheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGlu 820
Db 2652 TTCAGCAATCATGGCCCTCTTGGTATGGGTAT---GGTGGCGGTCTGAATAATTTTGGAA 2708
Qy 821 ArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyr 840
Db 2709 AGATTTTCTACATCACTCCATCGTATCTCTGGACATCTATTCCCTCTTGGCTTAC 2768
Qy 841 CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr 860
Db 2769 TGTATACATGGCTGCCATCTGTTATTTACAGCGAAATTTATCACTCCAGAGCTGAATAAT 2828
Qy 861 PheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGlu 880
Db 2829 GTTGCAGCGCTGTGGTTCATGCTATCTGTTTATCTGATCTTTTGTACGAGCATCTCAGAA 2888
Qy 881 MetArgTrpSerGlyValSerIleGluGluTyrTrpArgAsnGluGlnPheTrpValIle 900
Db 2889 ATGAGATGGAGTGGTGGTGGATTTGATGCTGTGGAGGATGACAGATTCTGGGTCAAT 2948
Qy 901 GlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLysValLeuAlaGly 920
Db 2949 GGAGGTGTCTCACACCTCTTGTGTGTTCCAGGACTTCTCAAGTCTATGCTGT 3008
Qy 921 IleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspAspGluPheAla 940
Db 3009 GTTGATACAAGCTTCAACGTCATCAAG-----GGTGGAGATGATGAGGAGTTCTCA 3062
Qy 941 GluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProThrThrLeuLeuIleIle 960
Db 3063 GAGCTATATACATTCATATGACTACCTATTATGATACCTCTTACCACCTTGCTTCTATTG 3122
Qy 961 AsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTrp 980
Db 3123 AACTTCATTTGGTGGTGGCTGGCTTTCATATGCGATCAATAACGATATGAGTCATGG 3182
Qy 981 GlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPhe 1000
Db 3183 GCCTCCCTCTTTGGAGAGCTATCTTGCATTTTGGGTGATGTCATCTTTATCCCTTT 3242
Qy 1001 LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIleTrpSerIle 1020
Db 3243 CTCAAAGGTTTGGTGGAGGCAAAACAGACCAACAGATGTGATCTGTGGTCCATT 3302
Qy 1021 LeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThr 1040
Db 3303 CTGCTGGCTTCAATCTCTCGCTCTTGGGTTCGATTTGATTCCTTTCTTTCGGAAGAT 3362
Qy 1041 LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3363 GATGTCGCTCTTCTTGGAGAGTGTGGTTGGATTGC 3398

RESULT 15

US-10-160-719-37

; Sequence 37, Application US/10160719

; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 37
; LENGTH: 3969
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (144)...(3401)
US-10-160-719-37

Alignment Scores:

Pred. No.: 0 Length: 3969
Score: 4073.50 Matches: 757
Percent Similarity: 79.68% Conservative: 129
Best Local Similarity: 68.08% Mismatches: 139
Query Match: 72.12% Indels: 87
DB: 16 Gaps: 15

US-10-627-132-30 (1-1052) x US-10-160-719-37 (1-3969)

Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
Db 144 ATGGAGGCGAGCGCGGCTGTGGCTCCACACCGCACGAGCTCGTCTGTCATC 203
Qy 21 Arg-----GlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGlu 38
Db 204 CGCCGCGACGCGCATCCCGGCGGAGCGCGGAGAGAGAGCGGCGAGGTGTGCCAG 263
Qy 39 IleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGlu 58
Db 264 ATTTGGCGGACGACGTCGGCTTCCCGGCGGAGACCCCTTCTGTCGGCTGCACACGAG 323
Qy 59 CysGlyPheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsn 78
Db 324 TGCCTTCCCGCTCTGCGGCGACTGTCTACGATACGAGCGCGGAGCGCGGAGCGACGAGAAC 383
Qy 79 CysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAsp 98
Db 384 TGCCCCAGTCCAGACTCGATACAGCGCTCAAGGCTCCCAACGTCGACCGGTGAC 443
Qy 99 AspAspGluGluAspIleAspAspLeuGluPheAsnIleAspAspGluAsnGln 118
Db 444 GAGGAGGAGGACGCGTCGATGACTGGACAAACAGTTCACTGGGACGGCCATGACTCG 503
Qy 119 GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
Db 504 CAG-----TCTGTGGCGGAGTCCATGCTCTACGCG 533
Qy 139 ArgMetSerTyrGlyArgGlyProAspAspGlyAsnAsnThrProGlnIlePro 158
Db 534 CACATGAGTACGCGCGTGA-----GTCACCTAATGGCGGCCCAAGCTTTC 584
Qy 159 ProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGly 178
Db 585 CAGCTC-----AACCCCAATGTTCCACTC-----CTCACCACACGGG 620
Qy 179 TyrGlyHisGlyGluValSerSerLeuHis----- 189

Db 621 CAAATGGTGGATGACATCCACCGGAGCAGCGCTGGTGCCTTCTTTCATGGGTGGT 680
Qy |||||
190 -----LysArgIleHisProTyrProValSerGluProGly-----SerAla 203
Db 681 GGGGAAAGAGGATACATCCCTTCTTATGCGGATCCACGCTTACCTGTCACACCGAG 740
Qy |||||
204 LysTrpAspGluLysGlu-----ValSerTrpLysGlu 215
Db |||||
741 TCTATGAGCCATCCAGGATCTTGTGTCATATGGGTATGGTAGTGTGTCTTGGAGGAA 800
Qy |||||
216 ArgMetAspTrpLysSerGln-----GlyIleLeuGlyGly 229
Db |||||
801 CGGATGAGGATTCGAAGCAGACAGAGAGGATGCACACGCGGGAATGATGGTGGT 860
Qy |||||
230 GlyAlaAspProGluAspMetAspAlaValAlaLeuAsnAspGluAlaAArgGlnPro 249
Db |||||
861 GGTGATGATGGTGAC-----GATGCTGATCTACCACTTAATGATGAAGCAACAA 914
Qy |||||
250 LeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleVal 269
Db |||||
915 CTGTCCAGAAATTCACCTTCATCAAGCCAGATTAATCCATATAGGATGATATCAT 974
Qy |||||
270 ValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAsp 289
Db |||||
975 ATTCGGCTTGTGGTCTTCTTCTTCCACTACCGAGTGATCCGGTGAATGAT 1034
Qy |||||
290 AlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIle 309
Db |||||
1035 GCATTTGCTTGTGGCTCATATCTGTATCTGTGAAATCTGGTGTGGCATGCTTGGATT 1094
Qy |||||
310 LeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSer 329
Db |||||
1095 CTTGATCAATTCACAAAGTGTTCCTATTCAGAGAGAGACTTACCTAGACCGGCTGTCA 1154
Qy |||||
330 LeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSer 349
Db |||||
1155 CTGAGTTTCGACAGGAAGCCAGCATCTCAACTTGTCTCAATTTGATTTCTTGTCAGT 1214
Qy |||||
350 ThrValAspProLeuLysGluProProLeuValThrAlaAnThrValLeuSerIleLeu 369
Db |||||
1215 ACGGTGTGATCCCTTAAAGGAACTCTTGTGTCAACAAATACTGTCTATCTATCTCT 1274
Qy |||||
370 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 389
Db |||||
1275 TCGGTGGATATTCCTGTTGATAAGTTTCTTGTATGTTTCTGATGATGGTGTGCAATG 1334
Qy |||||
390 LeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCys 409
Db |||||
1335 CTAACGTTTGAAGCATATCTGAACATCTGAATTTGCAAGAAATGGTTCCTTCTGTC 1394
Qy |||||
410 LysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyr 429
Db |||||
1395 AAACGGTACAAATATTTGAACCTCGCGCTCCAGAGTGGTACTTCCAAACAGAAATAGACTAC 1454
Qy |||||
430 LeuLysAspLysValGlnProThrPheValGlnArgArgAlaMetLysArgGluTyr 449
Db |||||
1455 TTGAAGACAAAGTGGCGCAACCTTGTGGGAGAGAGAGCAATGAGAGAGAGAT 1514
Qy |||||
450 GluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGlu 469
Db |||||
1515 GAGGAATTCAGGTGAGAATCAATGCCTTAGTTGCAAGGCCAGCAAGTTCCTGAAGAA 1574
Qy |||||
470 GlyTrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly 489
Db |||||
1575 GGATGGACAAATGCAAGATGGAAACCCCTGGCCCTGGAAACAAATGTTCTGATCATCCTGGA 1634
Qy |||||
490 MetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGluAsnGluLeuPro 509
Db |||||
1635 ATGATTCAGTCTTCTTGGCCAAAGCGGAGGCGCTTGACTGTGAGGGGAAATGAATGCGCA 1694
Qy |||||
510 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisIleLysLysAlaGly 529
Db |||||

Db 1695 CGATGGTTTATGTTTCTAGAGAGAAACGACCGAGGCTATAACCATCATTAAGAAAGCTGGT 1754
Qy |||||
530 AlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsn 549
Db |||||
1755 GCTATGATGATGCTGGTCCGAGTCTCTGCTGTACTAACAATGCTCCATATTTGTTAAAC 1814
Qy |||||
550 LeuAspCysAspHisTrpIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeu 569
Db |||||
1815 TTGGATTGTGATCACTACATCAACAACAGCAGCTATAAAGGAAGCAATGTGTTTATG 1874
Qy |||||
570 MetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGly 589
Db |||||
1875 ATGAGCCCTTACTACGAAAGAGGTTTGTATGTACAGTTCCTCAAGATTTGATGGG 1934
Qy |||||
590 IleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLys 609
Db |||||
1935 ATTGATCGCATGACCGATATGTACCGGAATGTGTCTTTTGTATATCAACATGAA 1994
Qy |||||
610 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGln 629
Db |||||
1995 GGTGGATGGTATTCAGGGTCCAAATTTATGTTGTACTCGATGTGTATTTAGAGCGAG 2054
Qy |||||
630 AlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysAsp 649
Db |||||
2055 GCATTATATGTTATGATGCCCCCAAA---ACAAGAAGCCCATCAAGGACTTGCAC 2111
Qy |||||
650 -----CysCysProCysPheGlyArgLysLysArgLysHisAla 662
Db |||||
2112 TGCTGGCCCAAGTGGTGTCTTGTCTGTCTTGGCAATAGGAAGCAAAAGAGACT 2171
Qy |||||
663 LysAspGlyLeuProGlu-----GlyThrAlaAspMetGlyValAspSerAspLysGlu 680
Db |||||
2172 ACCAAACCCAAACACAGAGAAGAAAGTATTATTTTCAAGAAAGAGAACCAATCC 2231
Qy |||||
680 -----GlyThrAlaAspMetGlyValAspSerAspLysGlu 680
Db |||||
2232 CCTGCATPATCTCTGTGTGAAATTCAGAGCTGCTCCAGAGCTGAGATGAAGAGGCC 2291
Qy |||||
681 MetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThr 700
Db |||||
2292 GGTATGTAATCAACAAATAATTAGAAAGAAATTTGCCCAATCTCTCTGTTTTGTACA 2351
Qy |||||
701 SerThrLeuMetGluGluGlyValProProSerSerSerProAlaAlaLeuLeuLys 720
Db |||||
2352 TCCACACTTCTCGAGATGTTGGAAACCTTGAAGAGTGCAGCTCTCTCTCTTTTGA 2411
Qy |||||
721 GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeu 740
Db |||||
2412 GAAGCTATACATGTCATTAGTTGGTTATGAAGACAGACAGACTGGGGAAGAGAGAT 2471
Qy |||||
741 GlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg 760
Db |||||
2472 GGCTGGATCTATGATCAGTTACAGAGATATTTCTAATCTGGTTCAAGATGCTATGTCAT 2531
Qy |||||
761 GlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIle 780
Db |||||
2532 GGTGGCGGTCAATTTACTGCATACCTAAACGGGTGCATTCAAAGGTTCTGCACCTCTG 2591
Qy |||||
781 AsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePhe 800
Db |||||
2592 AATCTTTTCAGATCTCTTCCACAGGCTCTCGGTGGGCTCTTGGGTCTTATTTGAGATCTTC 2651
Qy |||||
801 PheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGlu 820
Db |||||
2652 TTGAGCAATCATTCGCCCTCTTGGTATGGGTAT---GGTGGCGGTCTGAATTTTGGAA 2708
Qy |||||
821 ArgPheAlaTrpIleAsnThrIleTyrProPheThrSerLeuProLeuLeuAlaTyr 840
Db |||||
2709 AGATTTTCTACATCAACTCCATCGTGATCTTGGACATCTATTCCCTCTTGGCTTAC 2768
Qy |||||
841 CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr 860
Db |||||
2769 TGTACATTGCGCATCTGTTTATTGACAGGGAATTTATCATCCAGAGCTGATTAAT 2828

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2005, 00:02:06 ; Search time 9650 Seconds
(without alignments)
17288.218 Million cell updates/sec

Title: US-10-627-132-29
Perfect score: 3443
Sequence: 1 ctgcgtgcctgcctgcga.....aaaaaaaaaaaaaaaaaaaaa 3443

Scoring table:  Gapop 60%0.5, Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 8

Total number of hits satisfying chosen parameters: 226460

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hcg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3443	100.0	3443	8	AY372246 Zea mays
2	83	2.4	3631	8	AK121170 Oryza sat
3	79	2.3	165909	8	AF005420 Oryza sat
4	79	2.3	187410	8	AP005579 Oryza sat
5	59	1.7	1255	8	AY483156 Hordeum v
6	50	1.5	3470	8	AY372244 Zea mays
7	49	1.4	814	5	BC058699 Danio rer
8	49	1.4	4919	9	BSM080426
9	48	1.4	1368	10	BC083654 Rattus no
10	48	1.4	1792	9	BC017482 Homo sapi
11	48	1.4	13981	9	AC141002 Homo sapi
12	47	1.4	431	6	CQ700614 Sequence
13	47	1.4	1409	10	BC058737 Mus muscu
14	47	1.4	1417	5	ZOPY18678
15	47	1.4	1431	9	BC039586 Homo sapi
16	47	1.4	1450	10	BC019205 Mus muscu
17	47	1.4	1897	10	BC048832 Mus muscu
18	47	1.4	2094	9	BC043552 Homo sapi
19	47	1.4	2513	10	BC018554 Human DNA

20	47	1.4	2902	9	HSU31501
21	47	1.4	3099	10	BC054839
22	47	1.4	170632	9	AL136321
23	46	1.3	467	11	BV047979
24	46	1.3	682	9	BC043523
25	46	1.3	1412	5	BC053171
26	46	1.3	1415	9	BC017176
27	46	1.3	1705	9	BC019066
28	46	1.3	2258	9	AK026494
29	46	1.3	2432	5	BC082653
30	46	1.3	2500	10	BC018294
31	46	1.3	2559	5	BC080919
32	46	1.3	3258	9	AB050269
33	46	1.3	5011	10	BC042459
34	46	1.3	61590	10	AC145269
35	46	1.3	119858	10	AC147157
36	46	1.3	120990	10	AC142225
37	46	1.3	145513	10	AC145351
38	46	1.3	146775	10	AC145552
39	46	1.3	149202	10	AC134863
40	46	1.3	149857	10	AC123677
41	46	1.3	150659	10	AC147159
42	46	1.3	152085	10	AC145299
43	46	1.3	155037	10	AC136452
44	46	1.3	156149	2	AC027556
45	46	1.3	166004	9	AL935036
46	46	1.3	168601	10	AC132276
47	46	1.3	169626	10	AC134562
48	46	1.3	171583	10	AC135805
49	46	1.3	171934	10	AC147461
50	46	1.3	177597	2	AC147118
51	46	1.3	177864	10	AC074315
52	46	1.3	181650	10	AC147716
53	46	1.3	183653	10	AC147253
54	46	1.3	185622	10	AC142269
55	46	1.3	187998	10	AC145515
56	46	1.3	189876	10	AC142215
57	46	1.3	197287	10	AC140405
58	46	1.3	202850	10	AC144894
59	46	1.3	231856	2	AC103140
60	46	1.3	233924	10	AC134579
61	46	1.3	269126	2	AC106218
62	46	1.3	313855	2	AC124866
63	45	1.3	750	10	BC064126
64	45	1.3	844	9	BC018449
65	45	1.3	925	10	BC023421
66	45	1.3	986	6	BD074339
67	45	1.3	1258	10	BC030462
68	45	1.3	1317	9	BSM0804695
69	45	1.3	1327	9	BC003651
70	45	1.3	1336	9	BC016051
71	45	1.3	1434	6	AX358790
72	45	1.3	1434	6	AX362283
73	45	1.3	1434	9	AY358965
74	45	1.3	1438	9	BC028166
75	45	1.3	1732	10	BC058743
76	45	1.3	1746	9	BSM0805351
77	45	1.3	1764	9	BC032010
78	45	1.3	1930	9	BC064417
79	45	1.3	1972	9	IR05ST046
80	45	1.3	1984	6	BD275102
81	45	1.3	2012	10	BC055062
82	45	1.3	2188	9	BC031566
83	45	1.3	2269	5	AB045997
84	45	1.3	2309	5	BC055978
85	45	1.3	2592	9	BC018986
86	45	1.3	3145	6	BD275095
87	45	1.3	3377	5	BC084504
88	45	1.3	4125	9	BC038588
89	45	1.3	4552	9	BC047935
90	45	1.3	114145	9	AL162407
91	45	1.3	115710	2	AL139015
92	45	1.3	116103	9	AL606465

U31501	Human fragi
BC054839	Mus muscu
AL136321	Human DNA
BV047979	S209P6432
BC043523	Homo sapi
BC053171	Danio rer
BC017176	Homo sapi
BC019066	Homo sapi
AK026494	Homo sapi
BC082653	Xenopus l
BC018294	Mus muscu
BC080919	Xenopus t
AB050269	Homo sapi
BC042459	Mus muscu
AC145269	Mus muscu
AC147157	Mus muscu
AC142225	Mus muscu
AC145351	Mus muscu
AC145552	Mus muscu
AC134863	Mus muscu
AC123677	Mus muscu
AC147159	Mus muscu
AC145299	Mus muscu
AC136452	Mus muscu
AC027556	Homo sapi
AL935036	Human DNA
AC132276	Mus muscu
AC134562	Mus muscu
AC135805	Mus muscu
AC147461	Mus muscu
AC147118	Mus muscu
AC074315	Mus muscu
AC147716	Mus muscu
AC147253	Mus muscu
AC142269	Mus muscu
AC145515	Mus muscu
AC142215	Mus muscu
AC140405	Mus muscu
AC144894	Mus muscu
AC103140	Rattus no
AC134579	Mus muscu
AC106218	Rattus no
AC124866	Rattus no
BC064126	Mus muscu
BC018449	Homo sapi
BC023421	Mus muscu
BD074339	50 human
BC030462	Mus muscu
BSM0804695	
BC003651	Homo sapi
BC016051	Homo sapi
AX358790	Sequence
AX362283	Sequence
AY358965	Homo sapi
BC028166	Homo sapi
BC058743	Mus muscu
AL834311	Homo sapi
BC032010	Homo sapi
BC064417	Homo sapi
AL355713	Homo sapi
BD275102	47 Human
BC055062	Mus muscu
BC031566	Homo sapi
AB045997	Macaca fa
BC055978	Xenopus l
BC018986	Homo sapi
BD275095	47 Human
BC084504	Xenopus t
BC038588	Homo sapi
BC047935	Homo sapi
AL162407	Human DNA
AL139015	Homo sapi
AL606465	Human DNA

c 93 1.3 132296 10 AL772283 AL772283 Mouse DNA
c 94 1.3 149567 10 AC136639 AC136639 Mus muscu
95 1.3 149994 10 AC147259 AC147259 Mus muscu
96 1.3 152944 10 AC144580 AC144580 Mus muscu
c 97 1.3 154136 2 AC113980 AC113980 Mus muscu
c 98 1.3 158469 10 AC145586 AC145586 Mus muscu
99 1.3 165604 9 AC008927 AC008927 Homo sapi
c 100 1.3 188444 10 AC150035 AC150035 Mus muscu

ALIGNMENTS

RESULT 1
AY372246 3443 bp mRNA linear PLN 01-SEP-2004
LOCUS Zea mays cellulose synthase catalytic subunit 12 (CesA12) mRNA,
DEFINITION complete cds.
ACCESSION AY372246
VERSION AY372246.1 GI:38532103
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3443)
AUTHORS Dhugga,K.S., Barreiro,R., Appenzeller,L., Wang,H., Niu,X.,
Carriagan,L. and Tomes,D.
TITLE Cellulose formation and its role in determining stalk strength in
maize

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3443)
AUTHORS Dhugga,K.S.
TITLE Direct Submission
JOURNAL Submitted (21-AUG-2003) Trait and Technology, Pioneer Hi-Bred
International, Inc., a DuPont company, 7300 NW 62nd Avenue,
Johnston, IA 50131, USA

FEATURES
Location/Qualifiers
1..3443
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
1..3443
/gene="CesA12"
/note="CESA12; ZmCESA12"
/codon_start=1
/product="cellulose synthase catalytic subunit 12"
/protein_id="AA23312.1"
/db_xref="GI:38532104"
/translations="MEASAGLVAGSHNRNELVIRGHDPKPLRALSGOVCEICGDEV
GLTVGDGLFVACNECPVPCRYEYEREGTQNCQCKTRKRLKGSFRVAGDDDEE
DIDDLHEFNIDENQRLGEMNQNSQITEAMLHGRMSYGRPGDDGNTPIQPIPI
ITGSRSPVSGEPFITTPYGHGEVSSSLHRIHPYVPSPFGSAKNDKEKESWKERMD
DKSKGILGGGADPEDMDALNDARQPLSRKVSIASSKVPNPRMVIWVRLVLA
FFLYRLHPVDAIGLWLSIICEIWFALISWLDQFPKFPFIDRBTLDRLSLRYER
EGPSLSLAVDLFVSTVDPLKEPPLVANTVLSILADVDPVKVSCVSDGASMLTF
ESLSSETAFARKVPPCKKFGIEPRAPFPYFSLKVDYLDKQVQPTFVQRRAMKREYE
EFKVRINALVAKAMKPAEGIMKQTPMPNTRDHPGMIQVFLGHSGHDTGNEGL
PLRVYVSREKRGVQFHKKAGAMNALIRVSLTNAPFNLNDCDRIINNSKAIREAM
CFMLDQVQRKVCYQVQPPQFDGIDVDRYANRNTVFQDINMKLGGDIQGPVVTGTC
VFRQALYGNVPKPRKPMVTCDCPCGPKRKRKHAQGLPEGTADMGVSDKEML
MSHNPKEKRGQSAAFVTSITMEEGVPPSSPAALLKEAHIVISCGVEDKTDWGLEL
GWLYGSITEDLLTGFKMHCGRWSVYCMKRAAPKGSAPINLSLDRLNQLVRWALGSVE
IFFSRSLPFIAGFNKLNKWLERFAYINTIITYPTSLPLAYCTLPVACLLTQKFLMP
SISTFASLFPIAFKMSIFATGILMEKRWSGVSBEMWNEQFWIGVSAHLFAVWGL
LKVLAGIDNFTVTSKATGDEDEDEFAELAFKWTLLIIPPTILLIINIVGVVAGISDA
INNGYQSGPLFKGLFPFAMVVIHLVYFPLKGLMGRNRTPTVVVWINSILLASIFSLW
VRIDPFIVRTKGPDPVQGCING"

ORIGIN

Query Match 100.0%; Score 3443; DB 8; Length 3443;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGCGCTCGCCCTCGCTCGCAATCGCGAATCTGTCGAGCACCTGAGGGTGGAGGCGCGAG 60
DB 1 CTGCGTGCCTCGCTCGCAATCGCGAATCTGTCGAGCACCTGAGGGTGGAGGCGCGAG 60
QY 61 AGTAGCTAGCAGCGCGCGCTCCGCGCGCATGAGGCGCAGCGCGGGCTGCTGCGCCGG 120
DB 61 AGTAGCTAGCAGCGCGCGCTCCGCGCGCATGAGGCGCAGCGCGGGCTGCTGCGCCGG 120
QY 121 CTCGCACAAACCGGAAACGAGCTGCTGATCCGCGGCGCACAGGAGACCCCAAGCGCTGCG 180
DB 121 CTCGCACAAACCGGAAACGAGCTGCTGATCCGCGGCGCACAGGAGACCCCAAGCGCTGCG 180
QY 181 GGGCTGAGCGGGCAGGTGTCGAGATATGCGCGCAGGAGTGGGCTCACGGTGCACCG 240
DB 181 GGGCTGAGCGGGCAGGTGTCGAGATATGCGCGCAGGAGTGGGCTCACGGTGCACCG 240
QY 241 CGACCTCTTCTGCTCGCTGCAACGAGTGCCTTCCCGCTGTGCCGCTCTGTCTACGAGTA 300
DB 241 CGACCTCTTCTGCTCGCTGCAACGAGTGCCTTCCCGCTGTGCCGCTCTGTCTACGAGTA 300
QY 301 CGAGCGCCGGGAGGCGACGCAAACTGCCCCAGTGCAAGACGCGCTACAAGCGCTCAA 360
DB 301 CGAGCGCCGGGAGGCGACGCAAACTGCCCCAGTGCAAGACGCGCTACAAGCGCTCAA 360
QY 361 GGGGAGCGCGAGGTGTCGCGGCGCATGACGAGGAGGACATCGACGACTGGAGCACGA 420
DB 361 GGGGAGCGCGAGGTGTCGCGGCGCATGACGAGGAGGACATCGACGACTGGAGCACGA 420
QY 421 GTTCAACATCGACGACGAGATCAGCAGAGGAGCTGGAGGGCAACATCGACGAACAGCCA 480
DB 421 GTTCAACATCGACGACGAGATCAGCAGAGGAGCTGGAGGGCAACATCGACGAACAGCCA 480
QY 481 GATACCGAGGCGATGCTCGACGGCAGATGAGTACGGGAGGGGCGCCGACGACGGCGGA 540
DB 481 GATACCGAGGCGATGCTCGACGGCAGATGAGTACGGGAGGGGCGCCGACGACGGCGGA 540
QY 541 CGGCAACACACCCCGCAGATCCCGCCCATCATCACCGGCTCCCGCTCCGTCGGCTGAG 600
DB 541 CGGCAACACACCCCGCAGATCCCGCCCATCATCACCGGCTCCCGCTCCGTCGGCTGAG 600
QY 601 CGGTGAGTTTCCGATTACCAACGGGTATGCGCCACGCGAGGCTCTCTTCTTCCCTGCACAA 660
DB 601 CGGTGAGTTTCCGATTACCAACGGGTATGCGCCACGCGAGGCTCTCTTCTTCCCTGCACAA 660
QY 661 GCGCATCCATCCGTACCTGTGTGAGCCAGGAGTGCAGAGTGGGACGAGAAAGA 720
DB 661 GCGCATCCATCCGTACCTGTGTGAGCCAGGAGTGCAGAGTGGGACGAGAAAGA 720
QY 721 AGTGAGCTGGAAGGAGAGATGACGACCTGGAAGTCCAAAGCAGGGCATCTCGCGCGCGG 780
DB 721 AGTGAGCTGGAAGGAGAGATGACGACCTGGAAGTCCAAAGCAGGGCATCTCGCGCGCGG 780
QY 781 CGCCGATCCCGAAGACATGCAACGCGCTGCGCACTGGAACGAGCGAGGCGAGCGCGCT 840
DB 781 CGCCGATCCCGAAGACATGCAACGCGCTGCGCACTGGAACGAGCGAGGCGAGCGCGCT 840
QY 841 GTCGAGGAAGGTGTCGATCGCGTTCGAGCAAGGTGAACCCGCTACCGGATGGTGTGCTGGT 900
DB 841 GTCGAGGAAGGTGTCGATCGCGTTCGAGCAAGGTGAACCCGCTACCGGATGGTGTGCTGGT 900
QY 901 GGTCTCTGTTGCTCGCCTTCTTCTCCGCTACCGTATCTTCCACCCCGTCCCGGACGC 960
DB 901 GGTCTCTGTTGCTCGCCTTCTTCTCCGCTACCGTATCTTCCACCCCGTCCCGGACGC 960
QY 961 CATCGGCTGTGGCTCTCTCATCATCTGCGAGATCTGGTTCGCCATCTCTCTGGATCCT 1020
DB 961 CATCGGCTGTGGCTCTCTCATCATCTGCGAGATCTGGTTCGCCATCTCTCTGGATCCT 1020
QY 1021 CGACCATGTTCCCAAGTGGTTCGCCATCGACCGCGAGAGCTACCTCGACGCGCTCTCCCT 1080

Db 1021 CGACCAAGTCCCAAGTGGTTCCCAATCGACCGGAGACGTACCTCGACCGCTCTCCCT 1080
Qy 1081 CAGGTACGAGAGGAAAGGGAGCGCTCGCTGCTCTCGCGGTGACACTGTTGTTGAGCAC 1140
Db 1081 CAGGTACGAGAGGAAAGGGAGCGCTCGCTGCTCTCGCGGTGACACTGTTGTTGAGCAC 1140
Qy 1141 GGTGGACCGCTCAAGGAGCGCGCTGGTGTGACCGCAACACCGTGTCTCCATCTCCG 1200
Db 1141 GGTGGACCGCTCAAGGAGCGCGCTGGTGTGACCGCAACACCGTGTCTCCATCTCCG 1200
Qy 1201 CGTAGACTACCCCTGGACAAAGTCTCTCTGTACGTCTCCGACACCGCGGTGATGCT 1260
Db 1201 CGTAGACTACCCCTGGACAAAGTCTCTCTGTACGTCTCCGACACCGCGGTGATGCT 1260
Qy 1261 GACGTTTCAAGTTCGAGTTCGAGACCGCGGAGTTCGCGCGCAAGTGGGTGCTTCTGCAA 1320
Db 1261 GACGTTTCAAGTTCGAGTTCGAGACCGCGGAGTTCGCGCGCAAGTGGGTGCTTCTGCAA 1320
Qy 1321 GAAGTTCGGGATCGAGCCCGCGCGCGCGGAGTTCCTCTCTCGCTCAAGGTGCACTACCT 1380
Db 1321 GAAGTTCGGGATCGAGCCCGCGCGCGCGGAGTTCCTCTCTCGCTCAAGGTGCACTACCT 1380
Qy 1381 CAAGGACAAAGTGCAGCCCACTTCTGTGACGAGCGCGCGCCATGAAGAGAGATGA 1440
Db 1381 CAAGGACAAAGTGCAGCCCACTTCTGTGACGAGCGCGCGCCATGAAGAGAGATGA 1440
Qy 1441 GGAAGTTCGAGTTCGAGTTCGAGACCGCGGAGTTCGCGCGCAAGTGGGTGCTTCTGCAA 1500
Db 1441 GGAAGTTCGAGTTCGAGTTCGAGACCGCGGAGTTCGCGCGCAAGTGGGTGCTTCTGCAA 1500
Qy 1501 GTGATCATGAGGACGGACCGCGGTGGCGCGCGGAAACAACCGCGACACCGCGGCAT 1560
Db 1501 GTGATCATGAGGACGGACCGCGGTGGCGCGCGGAAACAACCGCGACACCGCGGCAT 1560
Qy 1561 GATCCAGGTGTTCTTGGGCAACAGCGCGCACGACACCGAGGCAACGAGCTGCCCCG 1620
Db 1561 GATCCAGGTGTTCTTGGGCAACAGCGCGCACGACACCGAGGCAACGAGCTGCCCCG 1620
Qy 1621 CTTCTGTGATCGTCTCCGTGAGAGCGCCCGGATTCAGCACCAAGAGGCGCGCGC 1680
Db 1621 CTTCTGTGATCGTCTCCGTGAGAGCGCCCGGATTCAGCACCAAGAGGCGCGCGC 1680
Qy 1681 CATGAAGCTCTGATTCGCTCTCGCGGTGCTGACCAACCGCGCATTCATGCTCAACTT 1740
Db 1681 CATGAAGCTCTGATTCGCTCTCGCGGTGCTGACCAACCGCGCATTCATGCTCAACTT 1740
Qy 1741 GGAAGTGTGATCACTACATCAACACAGCAAGGCCATCCGGAGGCGCATGCTTCTCAT 1800
Db 1741 GGAAGTGTGATCACTACATCAACACAGCAAGGCCATCCGGAGGCGCATGCTTCTCAT 1800
Qy 1801 GGAAGTGTGATCACTACATCAACACAGCAAGGCCATCCGGAGGCGCATGCTTCTCAT 1860
Db 1801 GGAAGTGTGATCACTACATCAACACAGCAAGGCCATCCGGAGGCGCATGCTTCTCAT 1860
Qy 1861 GGAAGTGTGATCACTACATCAACACAGCAAGGCCATCCGGAGGCGCATGCTTCTCAT 1920
Db 1861 GGAAGTGTGATCACTACATCAACACAGCAAGGCCATCCGGAGGCGCATGCTTCTCAT 1920
Qy 1921 GCTGGAGGATCCAGGCGCGGTGATGCTGCGGACAGGCTGCTGCTGCGGCGCGAGG 1980
Db 1921 GCTGGAGGATCCAGGCGCGGTGATGCTGCGGACAGGCTGCTGCTGCGGCGCGAGG 1980
Qy 1981 GCTCTAGGCTCAACACCTCCAGGAGGACCAAGAGGCCCAAGATGGTGAACCTGCGACTG 2040
Db 1981 GCTCTAGGCTCAACACCTCCAGGAGGACCAAGAGGCCCAAGATGGTGAACCTGCGACTG 2040
Qy 2041 CTGCGCGTGTCTGCGCGCAAGAGCGGAAACAGCGCAAGGACGGGTGCGCGAGGCGAC 2100
Db 2041 CTGCGCGTGTCTGCGCGCAAGAGCGGAAACAGCGCAAGGACGGGTGCGCGAGGCGAC 2100
Qy 2101 CGCTGATATCGGAGTACATAGCGACAGGAGATGCTCATCTCCACATGAATTCGAGAA 2160
Db 2101 CGCTGATATCGGAGTACATAGCGACAGGAGATGCTCATCTCCACATGAATTCGAGAA 2160

Db 2101 CGCTGATATCGGAGTACATAGCGACAGGAGATGCTCATCTCCACATGAATTCGAGAA 2160
Qy 2161 GCGGTTTCGGGACAGTTCGCGCGGTTCGTCAGTCAGTCGCTGATGAGAGAAAGCGCGCTCC 2220
Db 2161 GCGGTTTCGGGACAGTTCGCGCGGTTCGTCAGTCAGTCGCTGATGAGAGAAAGCGCGCTCC 2220
Qy 2221 TCCGTTTCGAGAGCCCGCGCGGTTCCTCAAGGAGGCCATCCATGTCTATGCTGCGGCTA 2280
Db 2221 TCCGTTTCGAGAGCCCGCGCGGTTCCTCAAGGAGGCCATCCATGTCTATGCTGCGGCTA 2280
Qy 2281 CGAGGACAAAGACCGACTTGGGGGCTGGAGCTGGGGTGGATCTAGCGGTTCGATCAGCGAGGA 2340
Db 2281 CGAGGACAAAGACCGACTTGGGGGCTGGAGCTGGGGTGGATCTAGCGGTTCGATCAGCGAGGA 2340
Qy 2341 CATCTTCGAGCGGGTTCGAAGATGCACTGCGCGGGTGGCGCTCGGTGTATGTCATGCGCGAA 2400
Db 2341 CATCTTCGAGCGGGTTCGAAGATGCACTGCGCGGGTGGCGCTCGGTGTATGTCATGCGCGAA 2400
Qy 2401 GCGGGCGGCGTTCGAAGGAGTTCGCGCGCGATCAATCTATCGGACCGGTCTCAACCGAGTGT 2460
Db 2401 GCGGGCGGCGTTCGAAGGAGTTCGCGCGCGATCAATCTATCGGACCGGTCTCAACCGAGTGT 2460
Qy 2461 CCGGTTCGCGCTGGGGTTCGCTCGAGATCTTCTTTCAGCGGACAGCGCCCTGCTGTACGG 2520
Db 2461 CCGGTTCGCGCTGGGGTTCGCTCGAGATCTTCTTTCAGCGGACAGCGCCCTGCTGTACGG 2520
Qy 2521 CTACAGAGACGGCAACCTCAAGTGGTGGAGCGCTTCGCGCTATCATCAACACCACTCTA 2580
Db 2521 CTACAGAGACGGCAACCTCAAGTGGTGGAGCGCTTCGCGCTATCATCAACACCACTCTA 2580
Qy 2581 CCGGTTTCACTTCGCTTCGCTTCGCTTACTGACCGCTTCGCGCGCTTCGCTTCCTCAC 2640
Db 2581 CCGGTTTCACTTCGCTTCGCTTCGCTTACTGACCGCTTCGCGCGCTTCGCTTCCTCAC 2640
Qy 2641 CGGCAAGTTCATCATGCGCGCTGATAGCAGTTCGCGCGCTCTTCTTTCATCGCGCTCTT 2700
Db 2641 CGGCAAGTTCATCATGCGCGCTGATAGCAGTTCGCGCGCTCTTCTTTCATCGCGCTCTT 2700
Qy 2701 CATGTCATCTTCGCGACCGGATTCCTGAGATGCGGTGAGCGGGTGGAGCATFCGAGGA 2760
Db 2701 CATGTCATCTTCGCGACCGGATTCCTGAGATGCGGTGAGCGGGTGGAGCATFCGAGGA 2760
Qy 2761 GTGGTGGAGGACGAGCAGTTCGCGGTTCATCGCGGGGCTGCGCGCATCTCTTCGCGGT 2820
Db 2761 GTGGTGGAGGACGAGCAGTTCGCGGTTCATCGCGGGGCTGCGCGCATCTCTTCGCGGT 2820
Qy 2821 CGTGAGGCGCTGCTCAAGGTTCCTCGCGGATTCGACCAACTTCACCGCTCACTCTCAA 2880
Db 2821 CGTGAGGCGCTGCTCAAGGTTCCTCGCGGATTCGACCAACTTCACCGCTCACTCTCAA 2880
Qy 2881 GGCACCGGCGACGAGGACGAGTTCGCGGAGCTTACGCGCTTCAAGTGGACACGCT 2940
Db 2881 GGCACCGGCGACGAGGACGAGTTCGCGGAGCTTACGCGCTTCAAGTGGACACGCT 2940
Qy 2941 CTTCTATCCCGCCACACCGCTGCTCATATTAAGTTCATCGCGGTTCGCGCGCATCTC 3000
Db 2941 CTTCTATCCCGCCACACCGCTGCTCATATTAAGTTCATCGCGGTTCGCGCGCATCTC 3000
Qy 3001 CGAGCGCATCAACAAAGGCTTCGCGGCGCTTCTTTCGCGAAGCTCTTCTTCGCG 3060
Db 3001 CGAGCGCATCAACAAAGGCTTCGCGGCGCTTCTTTCGCGAAGCTCTTCTTCGCG 3060
Qy 3061 CTTTCTGGGTCATCGTCCACCTTACCGTTCCTCAAGGGGCTCATGGGGCGCGAGACAG 3120
Db 3061 CTTTCTGGGTCATCGTCCACCTTACCGTTCCTCAAGGGGCTCATGGGGCGCGAGACAG 3120
Qy 3121 GAGCGGACCGTTCGTCATCTGTCGCTTCTGTCGCTTCATCTTCTCCCTGCTCTG 3180
Db 3121 GAGCGGACCGTTCGTCATCTGTCGCTTCTGTCGCTTCATCTTCTCCCTGCTCTG 3180
Qy 3181 GGTTCAGGATCGACCGCTTTCATCTGTCAGGACCAAGGCGCGGAGCTCAGGCAAGTGTGCGAT 3240
Db 3181 GGTTCAGGATCGACCGCTTTCATCTGTCAGGACCAAGGCGCGGAGCTCAGGCAAGTGTGCGAT 3240

QY 3241 CAATTGCTGAGCTGTTTATTAAAGGTTCAAAATTTCTGGAGCTTGTGTCATAGGAGAAAAA 3300
 |||||
 Db 3241 CAATTGCTGAGCTGTTTATTAAAGGTTCAAAATTTCTGGAGCTTGTGTCATAGGAGAAAAA 3300
 |||||
 QY 3301 ACAATTTAGAAATTTTGAAGGTTGTGTGCTGTAATGTTATGGTACCAGAAATGTGCG 3360
 |||||
 Db 3301 ACAATTTAGAAATTTTGAAGGTTGTGTGCTGTAATGTTATGGTACCAGAAATGTGCG 3360
 |||||
 QY 3361 GACGAGGAATTCGACAAAGCAAGGTTTGAATGTTAAATGCGCAAAAAA 3420
 |||||
 Db 3361 GACGAGGAATTCGACAAAGCAAGGTTTGAATGTTAAATGCGCAAAAAA 3420
 |||||
 QY 3421 AAAAAA 3443
 |||||
 Db 3421 AAAAAA 3443
 |||||

RESULT 2
 AKI211170
 LOCUS
 DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J023081B08, full insert sequence.
 AKI211170 3631 bp mRNA linear PLN 29-OCT-2003
 Oryza sativa (japonica cultivar-group)
 ACCESSION
 VERSION AKI211170.1 GI:37990793
 FLI CDNA: CAP trapper.
 KEYWORDS Oryza sativa (japonica cultivar-group)
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
 1 The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ootomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurokawa, R., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariakawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Kaga, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.
 Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

TITLE
 JOURNAL Science 301 (5631), 376-379 (2003)
 MEDLINE 22752273
 PUBMED 12869764
 REFERENCE
 2
 AUTHORS
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, T., Nagata, T., Nakahama, Y., Nakamura, M., Namiki, T., Nariakawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akashira, S.,

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

URL : http://cdna01.dna.affrc.go.jp/cDNA/
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Yamamoto, M., and Nakahama, Y.
 PAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariakawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K., and Murakami, K.
 Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Tomaru, A., Waki, K., Takaku-Akashira, S., Tanaka, T., Tomaru, A., Toyama, Y., Yasunishi, A., and Hayashizaki, Y.
 Location/Qualifiers
 source
 1. 3631
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J023081B08"

ORIGIN

Query Match 2.4%; Score 83; DB 8; Length 3631;
 Best Local Similarity 100.0%; Pred. No. 5.8e-32;
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 813 CACTGAACGACGAGCGAGCGAGCGCTGTCGAGGAAGGTGTCGATCGCGTCGAGCAAGG 872
 Db 823 CACTGAACGACGAGCGAGCGAGCGCTGTCGAGGAAGGTGTCGATCGCGTCGAGCAAGG 882
 QY 873 TGAACCCGTTACCGGATGGTGATC 895
 Db 883 TGAACCCGTTACCGGATGGTGATC 905
 RESULT 3
 AP005420
 LOCUS
 DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 9, PAC clone:F041B08.
 ACCESSION AP005420
 VERSION AP005420.3 GI:50725883
 KEYWORDS

SOURCE	Oryza sativa (japonica cultivar-group)	
ORGANISM	Oryza sativa (japonica cultivar-group)	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
	Ehrhartoideae; Oryzaceae; Oryza.	
REFERENCE	1	
AUTHORS	Sasaki, T., Matsumoto, T. and Katayose, Y.	
TITLE	Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, PAC clone: P0418B08	
JOURNAL	Published Only in Database (2002)	
REFERENCE	2 (bases 1 to 165909)	
AUTHORS	Sasaki, T., Matsumoto, T. and Katayose, Y.	
TITLE	Direct Submission	
JOURNAL	Submitted (12-JUN-2002) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan	
	(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/, Tel: 81-298-38-7441, Fax: 81-298-38-7468)	
COMMENT	On Jul 27, 2004 this sequence version replaced gi:41152735. Genes were predicted from the integrated results of the following: GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH (http://www.softberry.com/), GeneMark.hmm (http://opal.biology.gatech.edu/GeneMark/), GlimmerM (http://www.tigr.org/db/glimmerm/glmr.form.html), RiceHMM (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP or DDBJ. Protein homologies of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence. The orientation of the sequence is from SP6 to T7 of the PAC clone. This sequence of P0418B08 clone has an overlap with OJ1740.D06 (DDBJ: AP005579) clone at 5' end and with OJ1299.All (DDBJ: AP005568) clone at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.	
FEATURES	Location/Qualifiers	
source	1..165909	
	/organism="Oryza sativa (japonica cultivar-group)"	
	/mol_type="genomic DNA"	
	/cultivar="Nipponbare"	
	/db_xref="taxon:39947"	
	/chromosome="9"	
	/clones="P0418B08"	
gene	complement(5396..5638)	
	/gene="P0418B08.1"	
misc_feature	complement(5396..5638)	
	/gene="P0418B08.1"	
	/note="hypothetical ORF predicted by GlimmerM"	
gene	5856..10431	
	/gene="P0418B08.2"	
mRNA	join(5856..5221,6327..6562,6658..6754,6846..6972,7059..7671,7761..8024,8107..8319,8442..8649,8741..8937,9019..9372,9476..10431)	
	/gene="P0418B08.2"	
	/note="supported by full-length cDNA(s): AK1211170"	
	join(5951..6221,6327..6562,6658..6754,6846..6972,7059..7671,7761..8024,8107..8319,8442..8649,8741..8937,9019..9372,9476..10063)	
	/gene="P0418B08.2"	
	/note="contains EST(s): C72208(E1220), AU082309(E1220)	
	contains full-length cDNA(s): AK1211170"	
	/codon_start=1	
	/product="putative cellulose synthase"	
	/protein_id="BAD33412.1"	
	/db_xref="GI:50725884"	
	/translation="MEASAGLVAGSHRNELVLIRGHEPKPLRALSGQVCEICGDEVGRVDDGLFVAGNECGFPVPCPCYEREGTQNCQCKTRYKRLGSGPRVPGDEEDIDDLHEFNIDDEKQKQDQDQWQNSHITAMLHGMYSYGRGDDGNSGTPPLPITGARSFVSGEFPISNGHGGEFSSSLHRIHPYFVSEPGSAKWEKKEVSKERM DWKSKQGVAGGADPDYDADPLNDARQPLSRKVSIASSKVNPKYRNVILRLVV LGFFLRYLTHVPDAIPWLTSICEIWFVSWIIDQPKWYPIIDRETVLDRLSRLY EREGESLISAVDLFVSTDPLKEPLVTANTVLSILAVDYPVDKVSVCYSDGASML ETEGSETAEFAKWWPFCCKFSIEPRAFEYFSQVDYLKOKHVFVERRMKRE YEEFKVINALVAKQKQVPAEGWIMKDGTFWPGNTRDHPGMIQVFLGHSGHDTGNE ELPLVYVSREKRPFGQHHKACAMNALLRVSALVTNAPMLNLDCHYILNNSKATRE AMCFMDPOGVKQVQVQFQRFQDGDVDRVANRVTFDINMKGLDQFGQVYVGT GCVFPEALYGVNPKGPKRBMWTCDCPCFCGRKRGKQGLPRAVADGWDSDK EMLMSQNFKEFGQSAAFVTSTLMEEGVPPSSSPAALLKEAIIHVISQCYEDTWG LBLGWIYGSITEDILTFGKWKHCRKSVYVCMKRAAFKGSAPINLSDRLNQVLRWALG SVEIFSRHSLLYGYKNGNLKWLRFSTYNTIYPTSLPILAYCTLPACVLLTGKF IMPTSTFASLFFIALFISATGILEMRWSGVSEEMWRNEQFWIGVSAHLFAVV QGLKVLAGIDNTFTVTSKATGDEDEFAELVAFKWTLLIPTTLLILNIIGVAGV SDAINNGSEAMGPLGKLPFAFWIVHLYPFLKGLMGRQNRQRTTIVIVNSVLLASIFS LWMRIDPTTIIKRGPDVRCQGINC"	
gene	complement(10866..17118)	
	/gene="P0418B08.3"	
mRNA	complement(join(10866..11328,11488..11595,12606..12695,12806..12922,13597..13716,14119..15432,16817..17118))	
	/gene="P0418B08.3"	
	/note="supported by full-length cDNA(s): AK100679"	
	complement(join(11263..11328,11488..11595,12606..12695,12806..12922,13597..13716,14119..15357))	
	/gene="P0418B08.3"	
	/note="contains full-length cDNA(s): AK100679"	
	/codon_start=1	
	/product="streptococcal hemagglutinin-like protein"	
	/protein_id="BAD33413.1"	
	/db_xref="GI:50725885"	
	/translation="MDVLKSDVKTKGILNETLRLPPLVPSEKHNASPVNRGDRVASRYK NGLSASATARTCTSPSGRTSANECTPEPKAQSADRRRSTPSSRVSTSTPAS RSVTPRVNTVTEGHKSRRITSTRNTDGLWPMANLSSFSQSESVVTPGNKDKVIPS GSDQTKQASVIAERKSPRLKKNIGECENAPQSDQFRRVIEQHRVPMQSGRVA SNILSIDMSDKAGRSVSTNISRGVSPKTLASEGTGKFNKSLDEVAARRLAIHAG GRDDKVSCHAYQSQTSKVSRSFRAVTLFVPLVHRSSSPKASSTVTSISRSFQ SPSRTPSPSRQSAGSIQSGVASPIISVMVDKKNKSNQIENIHQLHLSYNYRL QMIFVNAVADTMSFOKVTAEIIYNNVLRNTSLRDVNMRRIMVOCIQQLKHLGTL KEQIDYLEQWPALEKENSISLPHATEALKASTLRLPVTSKADVVALKAVASAVDI KEGISAVRCMLPKVEDRITLVSELSVIARQEKMLDECRELLAMAKLQVQSSSLRT HQTGSRPGIAHMI"	
gene	18866..19216	
	/gene="P0418B08.4"	
misc_feature	18866..19216	
	/gene="P0418B08.4"	
	/note="hypothetical ORF predicted by GlimmerM"	
	this category is not included in IRGSP standard"	
gene	25510..27251	
	/gene="P0418B08.5"	
misc_feature	25510..27251	
	/gene="P0418B08.5"	
	/note="probably inactive due to including stop codon(s) in CDS"	
	pseudogene, transposase"	
	/pseudo	
gene	complement(29005..36438)	
	/gene="P0418B08.6"	

nrna	complement(join(<29005..29109,29221..29352,29652..29818,30448..30637,31719..31850,32464..32571,32702..32797,32891..32998,33106..33278,34215..34590,35076..35219,35314..35988,36106..36171,36271..36438)) /gene="P0418808.6" /note="start and end point are not identified"	REFERENCE AUTHORS TITLE	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae; Oryza.
CDS	complement(join(29005..29109,29221..29352,29652..29818,30448..30637,31719..31850,32464..32571,32702..32797,32891..32998,33106..33278,34215..34590,35076..35219,35314..35988,36106..36171,36271..36438)) /gene="P0418808.6" /note="similar to Oryza sativa chromosome 7, OSJNBa0061L20.19" /codon_start=1 /product="hypothetical protein" /protein_id="BAD33414.1" /db_xref="GI:50725886" /translation="MASDRISDDECCSNKSDNKVADSYANSISEEETSDSKYI TKKDLVLTLDIFSISIKRKRARTDKPVRQKRTKLEYDQDQSGSAEPPKRSRLN ITYFNLIEGLSEQRISIENAFGLNLFQRCALPLSFVKWIASHTDVSCSDIVVNG RSIPINPTNFIILGIPNGLEIKNDNDAGKHFFHQHFGSTKPLISFFGTLLSDKGV NKLSDDDLRCPMVALSTFLCPNSDTHPSKVLRLPLSDIKSSXWNNSKFVYEWLMT YIAKQKSKSEQTKLTGGCHLLAINLYLPKDFGRIENIPGPRISVWKGGMKIK EYKIDDECKTGFGRPLIDISTVCKMFTREHSDTTTGGNQLAAKLVDCVSKAS LFTQKDTTTRBEKSGEKSQGGNNVAHNPQLEQHNKNTDVCSQLPKTP QTKTAEHKTSINNTSISKAPPTDKKASQDFVCSQSNNSVASRTRAKRLHSAS PLSVVISPKKYSQDKIRKAVRKQPKATEPINTQPDRESNKDALFVEICTIPA KKEVQPTNLESSTQFVIDIEGPDADITGHTTKTKTFLVNSNSRHSNSQDP TODESDNIPKSTNPKFECLEHFDKSKTKNSASIGHFSLTSTGTNIMDIEHSMQ SQSQSPDIGNSPRIAQMRPNQHAQERYSMIRIDSLNASNCSTGRINLRPK RVHPLPCLNDYVWSSLSISLPRGVVDVYLNAYAKKIANDQNNKENEYRNFYFFH RTSLFFPCLNDYVWSSLSISLPRGVVDVYLNAYAKKIANDQNNKENEYRNFYFFH STVENKFSKFDIQYPPITRQNG" 40242..41638 /gene="P0418808.7" join(<40242..40308,40655..40900,40996..41112,41239..41339,41570..41638) /note="start and end point are not identified" /gene="P0418808.7" join(40242..40308,40655..40900,40996..41112,41239..41339,41570..41638) /gene="P0418808.7" /note="predicted by GeneMark.hmm etc." /codon_start=1 /product="hypothetical protein" /protein_id="BAD33415.1" /db_xref="GI:50725887" /translation="WPATELEEATATWELDEAGKANGGALPFPHPKPTAAVTRR	JOURNAL REFERENCE AUTHORS TITLE JOURNAL	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502 1503 1504 1505 1506 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544 1545 1546 1547 1548 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558 1559 1560 1561 1562 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572 1573 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660 1661 1662 1663 1664 1665 1666 1667 1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1

gene
FAGKQLEDGRTADYNIOKESTLHLVLRGGTMIKVKTLTGKEIIDEIPTDTIDRI
KERVEKGIIPVQQRILYAGKQLADDTKADYNIEGSGVHLVLRGSGY"
complement (2464. .8124)
/gene="OJ1740_D06.2"
mRNA
complement (join(2464. .3329,3566. .4732,5547. .5616,
7723. .8124))
/gene="OJ1740_D06.2"
/notes="supported by full-length cDNA (s): AK065475"
complement (join(2772. .3329,3566. .4732,5547. .5616,
7723. .7784))
CDS
/gene="OJ1740_D06.2"
/notes="contains EST(s): AU092173(C11159), D22463(C11159)
contains full-length cDNA(s): AK065475"
/codon_start=1
/product="putative RPT2"
/protein_id="BAD33627.1"
/db_xref="GI:50726106"
misc_feature
/translation="MKFMKLGSKPDAPQSDGADRVVVISPLATDVIHVSEVKFYLHK
FPILLSKSLQRLVIRKATGTEDEVIDGPGVTAPEICAKFCYGMVTLSPHNVA
ARCAAYLEMTEDVGNLFIKIDVFNSSILSKWDKIIIVLSTKALUPWSELKVI
GRCDIAIAKTSVDPANVTSYSHSRKGMCTEIVESTGRTSIAPKDWVYDLCELD
DLKRVVAVKSGRSPELIGALKAYAVRPLDSDYDALVAEDYMRNOCIVETIIV
LLPSDKTSGCSGFLKLLKVAITLVAGOHVKEELMERISFQLHKASVKDLGLPAASP
SDGADHVKLVNLRVQVARTAMSHNGGFVEKSDDKMIELNFQESTLALGELYDGYL
SEVASPDLSTFVELATVAPEARPHVDSISAVDAYLKEHFNISKADKKIKCGLI
DVKKLSTDSAMHATQNDRLPLRVQVLFQQLRAGSSNNAITDGGGHTCAKPIKMD
QSDICERRIPRHNSLNKQATSLSAREVHRKSGHRGSRNSFKDQGLLQSRSR
IFDKIWSKQGGKSGSETSGSPPLSAKPADVKPSPPLPLRNRYSVS"
complement (join(10055. .10068,13432. .13638,13766. .13899))
/gene="OJ1740_D06.3"
misc_feature
complement (join(10055. .10068,13432. .13638,13766. .13899))
/gene="OJ1740_D06.3"
/notes="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
gene
14864. .15751
/gene="OJ1740_D06.4"
misc_feature
14864. .15751
/gene="OJ1740_D06.4"
/notes="probably inactive due to 5' exon missing in CDS
pseudogene, putative HCWP repeat containing protein"
pseudo
16561. .17118
/gene="OJ1740_D06.5"
misc_feature
16561. .17118
/gene="OJ1740_D06.5"
/notes="probably inactive due to 5' exon missing in CDS
pseudogene, putative HCWP repeat containing protein"
pseudo
complement (17924. .18233)
/gene="OJ1740_D06.6"
mRNA
complement (join(<17924. .17955,18104. .>18233))
/gene="OJ1740_D06.6"
/notes="start and end point are not identified"
complement (join(17924. .17955,18104. .18233))
CDS
/gene="OJ1740_D06.6"
/notes="similar to Oryza sativa chromosome 1, P0480E02.21"
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD33628.1"
/db_xref="GI:50726107"
misc_feature
/translation="MKMKKIVDDDKRNRHDEEDVGDPAANSVATNDNDKARQHLHN
ASLGGELMK"
gene
join(18381. .18410,18539. .18577,19649. .19732)
/gene="OJ1740_D06.7"
misc_feature
join(18381. .18410,18539. .18577,19649. .19732)
/gene="OJ1740_D06.7"
/notes="hypothetical ORF
predicted by GlimmerM
this category is not included in IRGSP standard"
gene
25476. .26366
/gene="OJ1740_D06.8"
misc_feature
25476. .26366

gene
/gene="OJ1740_D06.8"
/notes="probably inactive due to 5' exon missing in CDS
pseudogene, TNP2-like transposon protein"
pseudo
26747. .28873
/gene="OJ1740_D06.9"
mRNA
complement (join(26747. .28873
27353. .28669
/gene="OJ1740_D06.9"
/notes="supported by full-length cDNA (s): AK065973"
misc_feature
/gene="OJ1740_D06.9"
/notes="putative Rim2 protein"
29731. .32622
/gene="OJ1740_D06.10"
misc_feature
29731. .32622
/gene="OJ1740_D06.10"
/notes="probably inactive due to 5' exon missing in CDS
pseudogene, putative hydroxyproline-rich glycoprotein"
pseudo
join(36559. .36654,36950. .37116,37456. .37510,38073. .38244,
38373. .38573)
/gene="OJ1740_D06.11"
misc_feature
join(36559. .36654,36950. .37116,37456. .37510,38073. .38244,
38373. .38573)
/gene="OJ1740_D06.11"
/notes="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
gene
complement (41633. .42004)
/gene="OJ1740_D06.12"
mRNA
complement (<41633. .>42004)
/gene="OJ1740_D06.12"
/notes="start and end point are not identified"
complement (41633. .42004)
CDS
/gene="OJ1740_D06.12"
/notes="predicted by GlimmerM etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD33629.1"
/db_xref="GI:50726108"
misc_feature
/translation="MTDEWPRVVRVRAELAVTGRGRPHVAAVVSLSYSGNHWSREIDDA
LTLFPLGLSLSCSGSREARETARDDDHDDHCCSCFVRVAGARSVSVIRLSQASRLPLF
AKSSPNSMRMRTVARTYHY"
42745. .42927
/gene="OJ1740_D06.13"
gene
817 GAACGACGAGCGGAGGCGCGCTGTCGAGGAAGGTGTCGATCGCGTCGAGCAAGGTGAA 876
Db 121846 GAACGACGAGCGGAGGCGCGCTGTCGAGGAAGGTGTCGATCGCGTCGAGCAAGGTGAA 121905
Qy 877 CCCGTACCGGATCGTGATC 895
Db 121906 CCCGTACCGGATCGTGATC 121924
Query Match 2.3%; Score 79; DB 8; Length 187410;
Best Local Similarity 100.0%; Pred. No. 6.1e-30; Indels 0; Gaps 0;
Matches 79; Conservative 0; Mismatches 0;
Qy 817 GAACGACGAGCGGAGGCGCGCTGTCGAGGAAGGTGTCGATCGCGTCGAGCAAGGTGAA 876
Db 121846 GAACGACGAGCGGAGGCGCGCTGTCGAGGAAGGTGTCGATCGCGTCGAGCAAGGTGAA 121905
Qy 877 CCCGTACCGGATCGTGATC 895
Db 121906 CCCGTACCGGATCGTGATC 121924
RESULT 5
AY483156
LOCUS
DEFINITION
Hordeum vulgare putative cellulose synthase catalytic subunit
(CesA8) mRNA, partial cds.
ACCESSION
AY483156
VERSION
AY483156.1 GI:39726036
KEYWORDS
Hordeum vulgare
ORGANISM
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
REFERENCE
1 (bases 1 to 1255)

AUTHORS Burton,R.A., Shirley,N.J., King,B.J., Harvey,A.J. and Fincher,G.B.
TITLE The Cesa Gene Family of Barley. Quantitative Analysis of
JOURNAL Transcripts Reveals Two Groups of Co-Expressed Genes
PUBMED Plant Physiol. 134 (1), 224-236 (2004)
14701917
REFERENCE 2 (bases 1 to 1255)
AUTHORS Burton,R.A., Shirley,N.J., King,B.J., Harvey,A.J. and Fincher,G.B.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-2003) School of Agriculture & Wine, University of
Adelaide, PMB 1, Glen Osmond, SA 5064, Australia
FEATURES
source Location/Qualifiers
1. .1255 /organism="Hordeum vulgare"
/mol_type="mRNA"
/db_xref="taxon:4513"
gene 1. .1255 /genes="Cesa8"
CDS <1. .1090 /genes="Cesa8"
/codon_start=2
/product="putative cellulose synthase catalytic subunit"
/protein_id="AA029968.1"
/db_xref="GI:39726037"
/translations="GRVGSAAFTVTFMEGGVPPSSPAALLKEAIHIVISCGYEDKT
DMGLEIGWYGSITDILTGFRMCHGRWSIYCMPLAAPKGSAPINLSDRLNQVLRW
ALGSVRIFFSRHSPLLYGKGNLKWLERFAYINTIYPTSLPLLAYCTLPVAVCLLT
GKFIIMPISITFASLPIFISIFATGILELRWSGVSIEWCRNEQLWVGWVAHLFP
AVIQGLLVLGADTKFTVTSKATGYEDDEFAELYAFKWTLLIPOTLLVNLIGV
AGISDAINNGYQSWGLFKLPFAFWIVHLPFLKGFGRQNRFTIIVISVLLAS
MFSLWVRIDPFTVKAGPDVQKQGINC"

Query Match 1.7%; Score 59; DB 8; Length 1255;
Best Local Similarity 100.0%; Pred. No. 3.1e-19;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2243 CTCCTCAAGAGGCCATCCATGTCATCAGCTCGGCTACGAGCAAGACCGACTGGG 2301
|||||
Db 83 CTCCTCAAGAGGCCATCCATGTCATCAGCTCGGCTACGAGCAAGACCGACTGGG 141
|||||

RESULT 6
AY372244 3470 bp mRNA linear PLN 20-JAN-2004
LOCUS Zea mays cellulose synthase catalytic subunit 10 (Cesa10) mRNA,
DEFINITION complete cds.
ACCESSION AY372244
VERSION AY372244.1 GI:38532099
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 3470)
AUTHORS Dhugga,K.S., Barreiro,R., Appenzeller,L., Wang,H., Niu,X.,
Carrigan,L. and Tones,D.
TITLE Cellulose formation and its role in determining stalk strength in
maize
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3470)
AUTHORS Dhugga,K.S.
TITLE Direct Submission
JOURNAL Submitted (21-AUG-2003) Trait and Technology, Pioneer Hi-Bred
International, Inc., a DuPont company, 7300 NW 62nd Avenue,
Johnston, IA 50131, USA
FEATURES
source Location/Qualifiers
1. .3470 /organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
gene 1. .3470 /genes="Cesa10"

CDS
29. .3265
/gene="Cesa10"
/note="CESA10; ZmCESA10"
/codon_start=1
/product="cellulose synthase catalytic subunit 10"
/protein_id="AAR23310.1"
/db_xref="GI:38532100"
/translation="MDAGSVTGLAAGSHMRDELHVMAREEPNAKRVADVKTCRCV
AEVGTREDGQPPVCAECGPVCPVYERSEGTQCPQCNTRYKROKQCPRVSGD
DEEGEMDDDFEDEFPAKSPKPEVADVSYNGEHPACKWRTGQTLSSFTSGVAG
KDLEAREMEGSMWKDRIDKWKTKQEKRGKLNHDDSDDDDKNEDEYMLLEAARQL
MRKVPSPSMINPYRIVIVLRVLVLCFFKFRITTPATDAVPLWLASVIELWFAFSW
ILDQLPKWAPVTRYTLDRLALRYDREGEACRLSPIDFVFTVPLKEPIITANTVL
SILADYKQVPTFCVSDSGSMLEFDALSETAFARRWVFCCKFAVPAPPEFYS
OKIDYLDKQVPTFCVKERRAMKREYBEFKVRINALVAKQKPEEGWQDGPWPGN
NTRDHPGMIOVLNGQALDVEGHELPRIVYVSREKRCYGNHKKAGANALVRVSAV
LNAPFLINLDCDHYVNSKAVREACFMFLDPQLGKLCYVQFPQRFQDGDHRYAN
RVNVPDINMGDGIQGVYTCVFNQALYGDPPRKRPRMTCDPWSWCCC
CCFGGKRGKARKDKGDEEPRRLGLGFKRKSCKDLGGSVAGSKGGGLYKK
HQAFLKEIEERLEGYDELSRLMSQKSFKEKRFQSPVFIASLVDGGLPQGA
DPAALIKAIHIVISCGYBEKTEWKEIGWYGSVTEIDILTGFRMCHGRWSVYCTPR
PAFKGSAPINLSDRLHVLWALGSVEIFMSRCHCLPRVAYGRLKWLERFAYNTIY
PFTSLPLAYCTIPAVCLLTGKFIPTLNNLASINFIALFLSIATSVLEERSGYSI
EDWRNEQFVWIGVSAHLFAVQFGLKVGVDTSFTVTSKAGDEADAFGDLFLPK
WTLVLVPPPTLLIINNVGIVAGVDANNVGSWGLPFGKLPFFSFWIVHLPFLKGL
MGRQNRTPITIVVLSILLASIFSLVWVRIDPFIKAGPILKPCGVEC"

Query Match 1.5%; Score 50; DB 8; Length 3470;
Best Local Similarity 100.0%; Pred. No. 1.7e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2771 AACGAGCAGTTCTGGGTCAATCGCGCGCTGTCGGCGCATCTCTTCGCCGT 2820
|||||
Db 2786 AACGAGCAGTTCTGGGTCAATCGCGCGCTGTCGGCGCATCTCTTCGCCGT 2835
|||||

ORIGIN
Danio rerio chromosome 20 open reading frame 149, like, mRNA (cdna
clone MGC:73395 IMAGE:4787352), complete cds.
BC059699 814 bp mRNA linear VRT 28-SEP-2004
LOCUS BC059699
DEFINITION BC059699.1 GI:37590413
ACCESSION MGC.
VERSION BC059699.1
KEYWORDS
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 814)
REFERENCE
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,E.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Frange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Sánchez,A., Helton,E., Kettner,M., Madan,A., Rodriguez,S.,
Fahney,J., Heltan,E., Kettner,M., Madan,A., Young,A.C., Greenchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
JOURNAL PUBMED

REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 814)
Director MGC Project.
Submitted (01-OCT-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: Susan E. Brockerhoff, University of Washington
DNA Library Preparation: Susan E. Brockerhoff, University of
Washington
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 136 Row: j Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 41152387.

FEATURES

source
1..814
Location/Qualifiers
/organism="Danio rerio"
/mol_type="mRNA"
/strain="wild-type"
/db_xref="taxon:7955"
/clones="MGC:73395 IMAGE:4787352"
/tissue_type="Eye, adult retina, 1-2 year old, mixed sex"
/clone_lib="zebrafish adult retina cDNA"
/lab_host="DH10B"
/note="vector: pBluescript SK-"
1..814
/gene="c20orf1491"
/note="synonyms: fj62f01, zgc:73395"
/db_xref="LocusID:336303"
166..513
/gene="c20orf1491"
/codon_start=1
/product="Unknown (protein for MGC:73395)"
/protein_id="AAH59699.1"
/db_xref="GI:37590414"
/db_xref="LocusID:336303"
/translation="MAAIPASGLVATHYRRIGTSSSSCGSSSEYSGEIVPHHP
GLPKQDSGHWSFFFGKQNLPGMTVAEEAQKSGVSVTNGQVTCVAREMYMRQAS
ESSDGGKSEAGNS"

gene

CDS

ORIGIN

Query Match 1.4%; Score 49; DB 5; Length 814;
Best Local Similarity 100.0%; Pred. No. 6.2e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3395 TTAATGCGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|||||
Db 754 TTAATGCGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 802

RESULT 8

HSM808426
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

HSM808426
Homo sapiens mRNA; cDNA DKFZp686B20125 (from clone DKFZp686B20125).
BX648278
BX648278.1 GI:34367437
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
CONSRM
TITLE
JOURNAL
COMMENT

1 (bases 1 to 4919)
Bloeker, H., Boecher, M., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobo, G., Han, M. and Wiemann, S.
The German Human cDNA Consortium
Direct Submission
Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr. 1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GSF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686B20125) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at <http://mips.gsf.de/proj/cDNA/>.

FEATURES
source
1..4919
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686B20125"
/tissue_type="human uterus"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfilA + SfilB"
/dev_stage="adult"
4844

ORIGIN

Query Match 1.4%; Score 49; DB 9; Length 4919;
Best Local Similarity 100.0%; Pred. No. 5.6e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3395 TTAATGCGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|||||
Db 4836 TTAATGCGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4884

RESULT 9

BC083654
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BC083654
Rattus norvegicus cDNA clone MGC:94487 IMAGE:7135192, complete cds.
BC083654.1 GI:53733462
MGC.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

1 (bases 1 to 1368)
AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,
Diatchenko, L., Marusina, K., Parker, A.A., Rubin, G.M., Hong, L.,
Spatalen, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Carnetz, T.E., Brownstein, M.J., Ustin, T.B., Toshitake, S.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., Loquellano, N.A., Peters, G.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smaluk, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

/translation="MGKNSKLREVLQDLRENTETDHELOEWYKGLKDCPTGHLT
VDEPKLIYANFPYGDASKAEHVFRTDNGDTDFRFFIALSVTSRGLKQKXK
WAFSMDLNGYISRSEMLEIYQAIYKVMSSWKMPEDSTPEKRTDKIFRQMDTNN
DGKLSLEEFIRGAKSDPSIVRLQLQCPSSASQ"

ORIGIN
Query Match 1.4% Score 48; DB 9; Length 1792;
Best Local Similarity 100.0%; Pred. No. 2e-13;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3396 TAAATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443

Db 1736 TAAATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1783

RESULT 11

AC141002/c

LOCUS

DEFINITION Homo sapiens 3 BAC RP11-57206 (Roswell Park Cancer Institute Human

BAC Library) complete sequence.

AC141002

VERSION

AC141002.8

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 138981)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,P.R., Allen,C.,

Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,

Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,

Bouch,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,

Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,

Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,

Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,

Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,

Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,

Delaney,K., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,

Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,

Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,

Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,

Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,

Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,

Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,

Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C.,

Hollins,B., Honsi,P., Howard,S., Huber,J., Hulyk,S., Hume,J.,

Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,

Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,

Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,D.C.,

Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,

Louleeged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,

Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,

Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,

Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,

Moore,S., Morgan,M., Morris,T., Morris,S., Moser,M., Neal,D.,

Nelson,D., Newton,N., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,

Nickerson,E., Nwokenko,S., Oguh,M., Okwuonu,G., Oragunye,N.,

Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,

Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,

Rojas,A., Rojebokan,I., Rolfe,M., Ruiz,S., Savary,G., Scherer,S.,

Scott,G., Shen,H., Shoshitari,N., Sisson,I., Sodergren,E.,

Sonaite,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,

Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,

Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,

Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,

Warren,R., Washington,C., Watlington,S., Williams,G.,

Williamson,A., Wleczky,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,

Xu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and

Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 138981)

Worley,K.C.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

COMMENT

Direct Submission
Submitted (07-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 138981)
Worley,K.C.

Direct Submission
Submitted (23-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 138981)
Worley,K.C.

Direct Submission
Submitted (26-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 138981)
Worley,K.C.

Direct Submission
Submitted (29-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Mar 26, 2003 this sequence version replaced gi:29164555.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
<http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht>
ml.

FEATURES

source

Location/Qualifiers
1. .138981

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="3"

/clone="RP11-57206"

/complement(144..215)

/rpt_family="AluYd3a1"

/complement(231..344)

/rpt_family="AluSq/x"

/587..701

/rpt_family="L2"

repeat_region

repeat_region

repeat_region

```
repeat_region complement(702..1016)
repeat_region /rpt_family="AluY"
repeat_region 1017..1172
repeat_region /rpt_family="L2"
repeat_region 1173..1412
repeat_region /rpt_family="L1"
repeat_region 1413..1644
repeat_region /rpt_family="AluSq"
repeat_region 1645..1795
repeat_region /rpt_family="L1"
repeat_region 1796..1938
repeat_region /rpt_family="AluSx"
repeat_region 1939..2239
repeat_region /rpt_family="AluY"
repeat_region 2240..2415
repeat_region /rpt_family="AluSx"
repeat_region 2416..2532
repeat_region /rpt_family="L1"
repeat_region 2548..2842
repeat_region /rpt_family="AluY"
repeat_region 2888..3108
repeat_region /rpt_family="AluJo"
repeat_region 3111..3212
repeat_region /rpt_family="L1"
repeat_region 3284..3426
repeat_region /rpt_family="L1"
repeat_region complement(3460..3767)
repeat_region /rpt_family="AluSc"
repeat_region complement(3831..3861)
repeat_region /rpt_family="Alu"
repeat_region 3870..3898
repeat_region /rpt_family="AT rich"
repeat_region complement(3900..4188)
repeat_region /rpt_family="AluJb"
repeat_region 4220..4336
repeat_region /rpt_family="LIME"
repeat_region 4360..4525
repeat_region /rpt_family="LIME"
repeat_region 4526..4817
repeat_region /rpt_family="AluSx"
repeat_region 4818..5134
repeat_region /rpt_family="LIME"
repeat_region 5136..5338
repeat_region /rpt_family="AluSx"
repeat_region complement(5411..5676)
repeat_region /rpt_family="AluJo"
repeat_region 6447..6752
repeat_region /rpt_family="AluSx"
repeat_region 6758..7062
repeat_region /rpt_family="AluSq"
STS 7187..7310
/standard_name="RH93866"
repeat_region 7479..7597
repeat_region /rpt_family="FLAM_A"
repeat_region 8012..8297
repeat_region /rpt_family="AluJb"
repeat_region 9581..9654
repeat_region /rpt_family="L2"
repeat_region 9755..10198
repeat_region /rpt_family="L2"
repeat_region 10220..10270
repeat_region /rpt_family="(TA)n"
repeat_region complement(10273..10569)
repeat_region /rpt_family="AluSx"
repeat_region 10743..11201
repeat_region /rpt_family="L2"
repeat_region complement(11202..11516)
repeat_region /rpt_family="AluJb"
repeat_region 11517..11527
repeat_region /rpt_family="L2"
repeat_region complement(11528..11800)
Query Match 1.4%; Score 48; DB 9; Length 138981;

Best Local Similarity 100.0%; Pred. No. 1.6e-13;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3396 TAAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 63816 TAAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 63769

RESULT 12
CQ700614 CQ700614 431 bp DNA linear PAT 03-FEB-2004
LOCUS Sequence 45540 from Patent WO02070737.
ACCESSION CQ700614
VERSION CQ700614.1 GI:42261381
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE Compositions and methods relating to osteoarthritis
JOURNAL Patent: WO 02070737-A 45540 12-SEP-2002;
ChondroGene Inc. (CA)
FEATURES
Location/Qualifiers
1..431
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 1.4%; Score 47; DB 6; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.4e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3397 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 337 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 383

RESULT 13
BC058737 BC058737 1409 bp mRNA linear ROD 30-JUN-2004
LOCUS Mus musculus RIKEN cDNA 1700020C11 gene, mRNA
DEFINITION IMAGE:5400435), complete cds.
ACCESSION BC058737
VERSION BC058737.1 GI:37589295
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Frange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettner,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,X.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
```

JOURNAL
PUBMED
12477932
REFERENCE
2 (bases 1 to 1409)
AUTHORS
Strausberg,R.
TITLE
Direct Submission
JOURNAL
Submitted (30-SEP-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland.
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayale,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Maiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Teurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAK Plate: 127 Row: f Column: 18
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 21312729.

FEATURES

source
Location/Qualifiers
1..1409
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="MGC:68085 IMAGE:5400435"
/tissue_type="Eye, retina, mouse strain C57Bl/6"
/clone_lib="NIH_MGC_94"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
gene
1..1409
/gene="1700020C11R1k"
/db_xref="LocusID:67900"
/db_xref="MGI:1916686"
100..600
/gene="1700020C11R1k"
/product="RIKEN CDNA 2610507A21"
/protein_id="AAH58737.1"
/db_xref="GI:37589296"
/db_xref="LocusID:67900"
/db_xref="MGI:1916686"
translations="MSEQRQGAERDLYRTWVYLGANEVGEAFRSIVPAAVWVLS
YGVSSVVLADAIKKGKAGEVPSPEAGNRMALAVVDTFVWQSLASVAIPGFTINR
LCAASLYVLGTWTHWPTVTKWTITLGLLAIPVILHPIDRSVDFLLDSSLRLKLYFSV
EKPS7P"

ORIGIN

Query Match 1.4%; Score 47; DB 10; Length 1409;
Best Local Similarity 100.0%; Pred. No. 6.9e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3397 AAATGGCAAAAAA 3443
|||||
Db 1340 AAATGGCAAAAAA 1386
|||||

RESULT 14
ZOPY18678

LOCUS
DEFINITION
Zosterisessor ophiocephalus mRNA for opsin.
ACCESSION
Y18678
VERSION
Y18678.1 GI:4210871
KEYWORDS
opsin.
SOURCE
Zosterisessor ophiocephalus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;
Gobioidi; Gobiidae; Zosterisessor.
REFERENCE
1 Archer,S.N. and Hirano,J.
AUTHORS
Comparative analysis of opsins in Mediterranean coastal fish
TITLE
Unpublished
JOURNAL
REFERENCE
2 (bases 1 to 1417)
AUTHORS
Archer,S.N.
TITLE
Direct Submission
JOURNAL
Submitted (25-JAN-1999) S.N. Archer, International Marine Centre,
Localita sa Mardini, 09072 Torregrande, Oristano, ITALY
FEATURES
source
Location/Qualifiers
1..1417
/organism="Zosterisessor ophiocephalus"
/mol_type="mRNA"
/db_xref="taxon:85428"
/tissue_type="retina"
/dev_stage="adult"
1..1059
/function="phototransduction"
/codon_start=1
/product="opsin"
/protein_id="CAA77260.1"
/db_xref="GI:4210872"
/db_xref="GOA:O9YGY9"
/db_xref="UniProt/Swiss-Prot:O9YGY9"
/translation="MNGTEGPPFIEMVNTGIVRSPYQYLYVNPAYACLGAYM
FFLLVFPVFNFTLLVLEHKKRLPTLLNLAVDLVDFVFGGFTTWYTSMRGT
FVLRLGNCIEGFFATLGGEIALSLVLAIERWVVVCKPISNFRGENHAIMGVAPT
WFMASACVPLVPLVWSRYIPEGQSCGDVYTRAGFNNESEFVIMFVHFCIPLAV
VGFCYGRLLCAVKEAAAAQOSETTORAEREVSRVWVIMVIGFLVCLWLPYASVAVYIF
THQGSFEGPFTVTVPFAFFAKSSSIYNPIYICNKKQFRHOMITTLCCGKNPFEDEGA
STYKTEASSVSSSSVSPA"
misc_feature
1..17
/note="degenerate primer"
ORIGIN
Query Match 1.4%; Score 47; DB 5; Length 1417;
Best Local Similarity 100.0%; Pred. No. 6.9e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3397 AAATGGCAAAAAA 3443
|||||
Db 1363 AAATGGCAAAAAA 1409
|||||
RESULT 15
BC039586
LOCUS
DEFINITION
Homo sapiens KIAA1704, mRNA (cdna clone MGC:48766 IMAGE:5722236),
complete cds.
ACCESSION
BC039586
VERSION
BC039586.1 GI:24660419
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1431)
AUTHORS
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Loquellano, N.A., Peters, G.J., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1431)
Strausberg, R.
Direct Submission
Submitted (01-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
Contact: anadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kerteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Series: IRAK Plate: 84 Row: i Column: 13
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 8923916.
Location/Qualifiers
1. .1431
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:48766 IMAGE:5722236"
/tissue_type="Ovary, pooled from 3 adults"
/clone_lib="NIH MGC 125"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

FEATURES
source

gene

1. .1431
/genes="KIAA1704"
/notes="synonyms: AD029, LSR7, ba245H20.2"
/db_xref="LocusID:55425"
104. .1126
/genes="KIAA1704"
/codon_start=1
/product="KIAA1704"
/protein_id="AAH39586.1"
/db_xref="GI:24660420"
/db_xref="LocusID:55425"

CDS

104. .1126
/genes="KIAA1704"
/product="KIAA1704"
/protein_id="AAH39586.1"
/db_xref="GI:24660420"
/db_xref="LocusID:55425"
/translation="MARDLIGPALPGFKVARGTAEDRPPVAGPALPPNVKSSSS
DSDSPEDSSSLVEQNGQESDEDGPTARKORKQDDDDDDGFGFALPPGFKKQ
DSDRPPIGALPGFIKSTQSKGRDDPGQETDSEDEDIIGMPAKGFVYNY
TTEFQRQRMKKLTKGDDSSKPIRESMTLPPENKQFGLGTRPRADDTSG
DISIWTDPADREKAKETQAKRSKSKDEEHILSRDLAEQVSSYNSKRSSEL
MDIHKHLKSKAAEDKNKQERIPFDRDKLVNRPDEAKKALIKKSRELNTRFSHG
KGNMPL"

ORIGIN

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Series: IRAK Plate: 39 Row: m Column: 2
This clone was selected for full length sequencing because it

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3397 AAATGCCAAAAA 1450 bp mRNA linear ROD 29-JUN-2004
||||| 3397 AAATGCCAAAAA 1450 bp mRNA linear ROD 29-JUN-2004
Db 1384 AAATGCCAAAAA 1450 bp mRNA linear ROD 29-JUN-2004
||||| 1384 AAATGCCAAAAA 1450 bp mRNA linear ROD 29-JUN-2004
BC019205 1450 bp mRNA linear ROD 29-JUN-2004
Mus musculus RIKEN CDNA 1700020C11 gene, mRNA (CDNA clone MGC:29077
IMAGE:5051472), complete cds.
BC019205
ACCESSION
VERSION
KEYWORDS
MGC.
BC019205.1 GI:17512505
Mus musculus (house mouse)
Mus musculus
Mus musculus

REFERENCE
AUTHORS

1 (bases 1 to 1450)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kerteman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1450)
Strausberg, R.
Direct Submission
Submitted (07-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nigri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Grant, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Scantrip, S., Thomas, P.J., Touchman, J.W., Tsurgou, C., Vogt, J.I., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

TITLE

human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1450)
Strausberg, R.

REFERENCE
AUTHORS

1 (bases 1 to 1450)
Strausberg, R.
Direct Submission
Submitted (07-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nigri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Grant, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Scantrip, S., Thomas, P.J., Touchman, J.W., Tsurgou, C., Vogt, J.I., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nigri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Grant, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Scantrip, S., Thomas, P.J., Touchman, J.W., Tsurgou, C., Vogt, J.I., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nigri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Grant, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Scantrip, S., Thomas, P.J., Touchman, J.W., Tsurgou, C., Vogt, J.I., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nigri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Grant, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Scantrip, S., Thomas, P.J., Touchman, J.W., Tsurgou, C., Vogt, J.I., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Series: IRAK Plate: 39 Row: m Column: 2
This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 6678975.

FEATURES source

Location/Qualifiers

1. .1450

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clones="MGC:23077 IMAGE:5051472"

/tissue_type="Liver, normal. 5 month old male mouse."

/clone_lib="NCI CGAP_Lig"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. .1450

/gene="1700020C11Rik"

/protein_id="AAH19205.1"

/db_xref="GI:17512506"

/db_xref="MGI:1916686"

160. .660

/codon_start=1

/product="RIKEN cDNA 2610507A21"

/protein_id="AAH19205.1"

/db_xref="GI:17512506"

/db_xref="MGI:1916686"

/translation="MSEQQRQGAERDLYRTWRYLGVANEGAFSLVPAVVWLS YGVSSVYLAIDDKGKAGVPSPEAGRNTRMALAVVDTFVWQSLSVAIPGFTINR LCAASLYLGTMTWPTVRKWTITLGLLAIPVIHPIDRSVDFDLSSLRKLYPSV EKPSTP"

ORIGIN

Query Match 1.4%; Score 47; DB 10; Length 1450;

Best Local Similarity 100.0%; Pred. No. 6.9e-13;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3397 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443

Db 1400 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1446

RESULT 17

BC048832

LOCUS

DEFINITION Mus musculus transducin-like enhancer of split 6, homolog of Drosophila E(spl), mRNA (cDNA clone MGC:55024 IMAGE:4506029), complete cds.

ACCESSION BC048832

VERSION BC048832.1 GI:28981430

KEYWORDS MGC.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1897)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schaefer, T.E., Brownstein, M.J., Udén, T.B., Toshiki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywicki, M.I., Skalek, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

TITLE

JOURNAL MEDLINE

22388257

PUBMED

12477932

REFERENCE

2 (bases 1 to 1897)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (14-MAR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgapsb-remail.nih.gov

Tissue Procurement: The Cepko Laboratory

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc_mgc@nhgri.nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 99 Row: e Column: 7.

Location/Qualifiers

1. .1897

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clones="MGC:55024 IMAGE:4506029"

/tissue_type="Eye, retina, mouse strain C57Bl/6"

/clone_lib="NIH_MGC_94"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. .1897

/gene="Tle6"

/note="synonyms: Grq6, 1810057E06Rik"

/db_xref="LocustID:114606"

/db_xref="MGI:2149593"

198. .1724

/gene="Tle6"

/codon_start=1

/product="Tle6 protein"

/protein_id="AAH48832.1"

/db_xref="GI:28981431"

/db_xref="LocustID:114606"

/db_xref="MGI:2149593"

/translation="MSIGVGAEGDSWPCVSHSHEAPGMSAQTENSKEEDKQVPESA LQHPKSTPGPOLPTRRRFLSDELDQPPVMDAEPQFCQGLIQLWELFMDSRQ KNOEGGEDSSQESKDSGLCDFKPEQPRHNSLSDADPFLIKSPSALLDYQEDV SRQPTQESSGCRADKFLKLSWGSEVLESSCNPSTALWQLERTVPAQLQKRVLK HQELLVAVSSFTRHVFTCSQGIKVMNLVNVQAEEDPESHUKSCVQDNKVVLRIC LLSSNRTLFPAGYVNLPGVITWDLAASPLSEKQCPGLSCQALANTKENMALGFT DGTVRIWDLRTQETVIRNLKPTNSARNLVKDDNIWTTGGLDACLRCKMRKAVLSLEH LFQSLMSLAHSPTEDELLGLANGQCLFNSRKRDQVLTVDTKDNTILGKFPSPNGK WVASVGNFNITVHSMPTGAKLFQVPEVGPVRCFDMTEGRLITGSRDCASVHIKY "

534. .1694

/gene="Tle6"

/note="COG2319; Region: FOG: WD40 repeat [General function prediction only]"

/db_xref="CDD:COG2319"

misc_feature

ORIGIN

```

Query Match      1.4%; Score 47; DB 10; Length 1897;
Best Local Similarity 100.0%; Pred. No. 6.8e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3397 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
    |||
Db 1848 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1894
    |||

RESULT 18
BC043552
LOCUS Homo sapiens, clone IMAGE:5174044, mRNA. PRI 13-JAN-2003
DEFINITION Homo sapiens, clone IMAGE:5174044, mRNA.
ACCESSION BC043552
VERSION BC043552.1 GI:27694536
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
MAMMALIA; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 2094)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: gcgbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
Info@bcsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabh, Parvaneh Saedi, Jacqueline
Schein, Duane Smalius, Michael Tsai, Lorraine Spence, Jeff Scott,
Michael Thorne, Miranada Tsai, Nataaja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 87 Row: 1 Column: 19.
FEATURES
source
Location/Qualifiers
1..2094
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5174044"
/tissue_type="Brain, Lung, Testis, adult, pooled whole"
/clone_lib="NIH MGC_115"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"

ORIGIN
Query Match      1.4%; Score 47; DB 9; Length 2094;
Best Local Similarity 100.0%; Pred. No. 6.7e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3397 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
    |||
Db 2031 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2077
    |||

RESULT 19
BC018554
LOCUS Mus musculus DNA segment, Chr 7, Brigham & Women's Genetics 0611
DEFINITION Mus musculus DNA segment, Chr 7, Brigham & Women's Genetics 0611

```

```

expressed, mRNA (cDNA clone MGC:19375 IMAGE:2646082), complete cds.
BC018554
VERSION BC018554.1 GI:17391320
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
EUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2513)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klauser, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Loquellano, N.A., Peters, G.J.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2513)
Strausberg, R.
Direct Submission
Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcgbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lounsbury, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 23 Row: a Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 26665872.
FEATURES
source
Location/Qualifiers
1..2513
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="MGC:19375 IMAGE:2646082"
/tissue_type="Mammary tumor metastasized to lung. Tumor
arose spontaneously from a senescent normal mammary
(clonal) outgrowth infected with the virus MMV."
/clone_lib="NCI CGAP_Lu29"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1..2513
/gene="DTBw0611e"

```

gene


```

/db_xref="LocusID:52857"
/db_xref="MGI:105490"
135. .2078
/genes="D7Bwg0611e"
/codon_start=1
/product="DNA segment, Chr 7, Brigham & Women's Genetics
0611 expressed"
/protein_id="AAH18554.1"
/db_xref="GI:17391321"
/db_xref="LocusID:52857"
/db_xref="MGI:105490"
/translation="MQSWSMLCPTYTKQRNEDFKRLFSKLPPEARLVISYCALQREI
LFGYSGARCDLFLFRWQNALLEKTLISQLEKTECLCKEKTKALPNAIOICTESKX
LFGYSGARCDLFLFRWQNALLEKTLISQLEKTECLCKEKTKALPNAIOICTESKX
QNLGLGCSPEVEGDMLASQITD SPGGAADSQSPSPVSGRGVTPNLASRASDDHAGNE
EDKEQETDQLDASSQSDTTPVPAELSPSEPTPDGPTSLGDLRLSRELLRTDTGNS
SSSTGEGDIALDLPLSGRLNLSVFMHGAERLQOMLFDSPSLQGLFQOKRFTDVT
LSPYSDSCKHQRVUTYITPISNQLGPKASVVEQTFLFRGPGAGCVSDSEVITQ
GIPWQDFYFATYRCTLGNARKNLSVSIKYRQPKVAKLSLIEKNSWGISGIEDYF
HJDLRLDAEAKELSEEGGQKTRGLLSGRKRPKPSLRWHRGDQGPDPDQCTQSTM
TDCSGLSRRFSPGAGCPGAGIPSAALVLSITVLITLINALLIFYRLWSLERTAHFT
EWSHLLALAKGFPDTPOTBESAILALQKHQFHSVEVHKWRQILTRASVELLDEMFKSLK
LHOGITVPPNPDPOTPODPPSP"

```

ORIGIN
 LMQGIIIVDFFLDIQFQDUSFF
 Query Match 1.4%; Score 47; DB 10; Length 2513;
 Best Local Similarity 100.0%; Pred. No. 6.7e-13;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3397 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|||||

Db 2437 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2483

RESULT 20				
HSU31501				
LOCUS	HSU31501	2902 bp	mRNA	linear
DEFINITION	Human fragile X mental retardation syndrome related protein (FXR2) mRNA, complete cds.			

ACCESSION	U31501.1	GI:1098636	
VERSION	U31501.1		
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		

REFERENCE	1 (bases 1 to 2902)
AUTHORS	Zhang, Y., O'Connor, J.P., Siomi, M.C., Srinivasan, S., Dutra, A., Nussbaum, R.L. and Dreyfuss, G.
TITLE	The fragile X mental retardation syndrome protein interacts with novel homologs FXR1 and FXR2
JOURNAL MEDLINE	EMBO J. 14 (21), 5358-5366 (1995)
PUBMED	96080171
REFERENCE	7489725
AUTHORS	2 (bases 1 to 2902)
TITLE	Zhang, Y. and Dreyfuss, G.
JOURNAL	Direct Submission
	Submitted (12-JUL-1995)
	Yan Zhang, Biochemistry and Biophysics, Howard Hughes Medical Institute, University of Pennsylvania, Clinical Research Building, Room 330, 422 Curie Boulevard, Philadelphia, PA 19104-6148, USA

```

/gene="FXR2"
/codon_start=1
/product="fragile X mental retardation syndrome related protein"
/protein=""
/protein_id="AAC30292.1"
/db_xref="GI:1098637"
/translation="MGGSLASGVDEPGLPVEVRGNSGAFYKGFVKDVHEDSVTIPFEN
NWOSERAPIPGDVLPLPPADNYKKEITGEDEVVSRANQEPSCGWLRAVRMMKGVDF
VIEVAGADAYANVILTVIRVAPNPNLPLATVGSFVAPEDDREACSNVWKEFV
KALGANCICFINIINSEFLITLSTEAPVKRASLGDMDHFRSLRTKLLMSRNEBATKH
LETSQALAAQAEFTWEDLWGLATGHGANIQARKVQKIQVEIVKSGVVRVVEGDNDK
ITEPAQCAQASYLEFSDSEVQVPRNLVGVKVGKNGKIQVEIVKSGVVRVVEGDNDK
KNPREGWVPFIWGTRENTISQAALLHYLSQVEVSLRLRLQLDQLRQIIGLGF
RPGSGRGSQKADGAYTSIDESSLHATRYGSGRGRRTGTGAPGPFSSDS
TASLSEKREPNRAGCPDPTGTGSERRRTTGGRGPPAPRTSRYNNSISIS
VKTQDPSNYSLLDTPESPEPPVSEPEGPVSAARRRRRRRTDORTVMDGGLS
DGNMTNLEDESRRKPNRRRRNRNNGTDDGISDGRDPTVADVIISRAESQSRQ
SAPLRTKTPESDLSLQKPGPSNGBLSAPLELGSWMVGVS"

```

ORIGIN

Query Match 1.4%; Score 47; DB 9; Length 2902;
Best Local Similarity 100.0%; Pred. No. 6.6e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3397 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|||||

Db 2843 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2889
|||||

RESULT 21	BC054839	LOCUS	BC054839	3099 bp	mRNA	linear	ROD 30-JUN-2004
DEFINITION	Mus musculus armadillo repeat containing, X-linked 2, mRNA (CDNA clone MGC:62562 IMAGE:6490133), complete cds.						

EC054839.1 GI:32766563
MGC.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3099)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altchul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Spateston,M., Soares,M.B., Donald,M.F., Casavant,T.L.,
Schettler,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaby,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywicki,M.I., Skalska,U., Smallos,D.E.,
Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932

gene	CDS
g1	g1
g2	g2
g3	g3
g4	g4
g5	g5
g6	g6
g7	g7
g8	g8
g9	g9
g10	g10
g11	g11
g12	g12
g13	g13
g14	g14
g15	g15
g16	g16
g17	g17
g18	g18
g19	g19
g20	g20
g21	g21
g22	g22
g23	g23
g24	g24
g25	g25
g26	g26
g27	g27
g28	g28
g29	g29
g30	g30
g31	g31
g32	g32
g33	g33
g34	g34
g35	g35
g36	g36
g37	g37
g38	g38
g39	g39
g40	g40
g41	g41
g42	g42
g43	g43
g44	g44
g45	g45
g46	g46
g47	g47
g48	g48
g49	g49
g50	g50
g51	g51
g52	g52
g53	g53
g54	g54
g55	g55
g56	g56
g57	g57
g58	g58
g59	g59
g60	g60
g61	g61
g62	g62
g63	g63
g64	g64
g65	g65
g66	g66
g67	g67
g68	g68
g69	g69
g70	g70
g71	g71
g72	g72
g73	g73
g74	g74
g75	g75
g76	g76
g77	g77
g78	g78
g79	g79
g80	g80
g81	g81
g82	g82
g83	g83
g84	g84
g85	g85
g86	g86
g87	g87
g88	g88
g89	g89
g90	g90
g91	g91
g92	g92
g93	g93
g94	g94
g95	g95
g96	g96
g97	g97
g98	g98
g99	g99
g100	g100

Email: cgapbs-x@mail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nhgri.nih.gov
 Akhtar, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghghi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 116 Row: j Column: 11
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 13385647.

FEATURES
 source
 1. .3099
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="MGC:62562 IMAGE:6490133"
 /tissue_type="Eye, retina, mouse strain C57Bl/6"
 /clone_lib="NIH_MGC_94"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 gene
 1. .3099
 /gene="Armcx2"
 /note="synonym: ALEX2"
 /db_xref="LocustID:67416"
 /db_xref="MGI:1914666"
 270. .2624
 /gene="Armcx2"
 /codon_start=1
 /product="armadillo repeat containing, X-linked 2"
 /protein_id="AAH54839.1"
 /db_xref="GI:32766564"
 /db_xref="LocustID:67416"
 /db_xref="MGI:1914666"
 /translation="MSRARDGCVAAIGVIGASAWCYKYTRGDKKKLTKPNR
 ASVTSGRAGLRAGFTIDLGFGFPNPVDIEIMNKAQGEASNLATVAEEVAPAA
 PSKVQNGAESKVOELNGAKTEANLESVMPSTCTVTPPKVAGLTAAEPIIGA
 PKYLEAPSTTEASGAVAAPGPTVSPMIAQTGPVPSPTTSTVSTGPAIIPVAHVGAV
 QSPGPAVPLPQVSLVPAAPSWAVAPPGAVYIPVAHFAGPAAASRVQSPGTVP
 LPPSSVLPKGVPSVPGRTVQSPCAVHPVAOSTGVVPPRAVOYSGAVTSGGA
 PSGAATPRAASTQRTASTAEVQVPRVAATATATPRIGTIPAMVAEASLPVHSGAA
 ENPTSGSKTAATKGAAPGAHTGAI PKAGSATGAVPVGKGGKGNKRSKGGKRN
 KXNVDELGMGRFPGDGAASAAAAANGGQAFIAEIPSEGESGWTDTSDSE
 DVPQKGGKRTIPMKRPFYDEILGVDRDLKVLALQKDDPFIQQVALI.TLGN
 ANYSCNOETIRKLGGLPIIANMINKTDPHI KEKALMANNLSNENQGLQVYMKV
 MDIMASNLNSAVQVGLKFLTNMTITNDYQHLLVNSIANFFLLSQGGGKIYEILK
 ILNSFAENPDLKLLGTQVPSFSSLYNSYSEILINALTIFELIFDNLRAEVNY
 REFNGSLFLYLTTSFGVCVKIKALANHHDLVKKVKLVNRP"

ORIGIN

Query Match 1.4%; Score 47; DB 10; Length 3099;
 Best Local Similarity 100.0%; Pred. No. 6.6e-13;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 3397 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
 Db 2999 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3045

RESULT 22

AL136321

LOCUS
 DEFINITION
 Human DNA sequence from clone RP11-31E23 on chromosome 1q31.3-32.1
 Contains STSS and GSSs, complete sequence.
 ACCESSION
 AL136321
 VERSION
 AL136321.5 GI:7801524
 KEYWORDS
 HTG.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 170632)
 Kimberley, A.
 DIRECT SUBMISSION
 TITLE
 Submitted (04-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
 requests: clonerequest@sanger.ac.uk
 COMMENT
 On May 14, 2000 this sequence version replaced gi:7798853.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
 on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
 Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr1>
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. RP11-31E23 is from
 the library RPCI-11.1 constructed by the group of Pieter de Jong.
 For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACE3.6
 This sequence is the entire insert of clone RP11-31E23 The true
 left end of clone RP11-16L9 is at 142476 in this sequence.
 FEATURES
 source
 1. .170632
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="q31.3-32.1"
 /clone="RP11-31E23"
 /clone_lib="RPCI-11.1"
 13. .565
 /note="match: GSS: Em:AQ479504"
 631. .699
 /note="23 copies 3 mer gga 76% conserved"
 1092. .1388
 /note="AluJb repeat: matches 1. .290 of consensus"
 3527. .3614
 /note="MIR repeat: matches 60. .151 of consensus"
 3617. .3689
 /note="MIR repeat: matches 180. .260 of consensus"
 3878. .4304
 /note="match: STS: Em:G61801
 match: GSS: Em:AQ284198"
 3894. .4030
 /note="MIR repeat: matches 7. .176 of consensus"
 4443. .4597
 /note="MIR repeat: matches 78. .238 of consensus"
 4765. .4921
 /note="MUTJ repeat: matches 1. .185 of consensus"

```
repeat_region 5272..5307
/Note="9 copies 4 mer acac 100% conserved"
repeat_region 6386..6551
/Note="TIGER1 repeat: matches 1..157 of consensus"
repeat_region 6541..6737
/Note="TIGER1 repeat: matches 2202..2418 of consensus"
repeat_region 6858..7424
/Note="L2 repeat: matches 804..1417 of consensus"
repeat_region 7780..8079
/Note="L2 repeat: matches 1840..2151 of consensus"
repeat_region 8244..8356
/Note="L2 repeat: matches 2385..2497 of consensus"
repeat_region 8519..8646
/Note="16 copies 8 mer aaaggaaa 63% conserved"
repeat_region 8836..9248
/Note="L1M4 repeat: matches 4826..5272 of consensus"
complement(9071..9799)
/Note="match: GSS: Em:AQ781273"
misc_feature 9071..9799
complement(9283..9777)
/Note="match: GSS: Em:AQ206508"
misc_feature 9283..9777
complement(9291..9795)
/Note="match: GSS: Em:AQ581541"
misc_feature 9291..9795
/Note="match: GSS: Em:B32946"
complement(9468..9799)
/Note="match: GSS: Em:AQ207860"
misc_feature 9468..9799
/Note="match: GSS: Em:AQ425540"
repeat_region 9815..10291
/Note="9 copies 4 mer gtgt 91% conserved"
repeat_region 9859..9894
/Note="9 copies 4 mer gtgt 91% conserved"
repeat_region 9862..9893
/Note="4 copies 8 mer tgtgtgt 93% conserved"
misc_feature 9898..10340
/Note="match: GSS: Em:AQ759351"
repeat_region 9910..10254
/Note="MER7A repeat: matches 1..346 of consensus"
repeat_region 10616..10904
/Note="AluSp repeat: matches 1..306 of consensus"
repeat_region 11166..11281
/Note="AluSp repeat: matches 1..306 of consensus"
repeat_region 11816..11930
/Note="58 copies 2 mer aa 71% conserved"
misc_feature 12033..12445
/Note="MLT1H repeat: matches 95..208 of consensus"
misc_feature 12036..12525
/Note="match: GSS: Em:AQ317924"
repeat_region 12840..12974
/Note="match: GSS: Em:AQ332817"
repeat_region 13131..13233
/Note="27 copies 5 mer aaatt 62% conserved"
misc_feature 13264..13665
/Note="L2 repeat: matches 2611..2710 of consensus"
match: STS: Em:B30545
/Note="match: GSS: Em:G37739"
repeat_region 13357..13660
/Note="MLT1J repeat: matches 1..308 of consensus"
repeat_region 13747..13884
/Note="L2 repeat: matches 2355..2496 of consensus"
repeat_region 13925..13983
/Note="L2 repeat: matches 2643..2709 of consensus"
repeat_region 15042..15272
/Note="WIR repeat: matches 1..252 of consensus"
repeat_region 17808..18584
/Note="MER21B repeat: matches 6..790 of consensus"
repeat_region 19043..19221
/Note="MER5A repeat: matches 1..187 of consensus"
repeat_region 19433..19512
/Note="10 copies 8 mer aaccaact 85% conserved"
misc_feature 20045..20493
/Note="match: GSS: Em:AQ463010"
repeat_region 23365..23728
/Note="THE1C repeat: matches 1..370 of consensus"
repeat_region 23729..25310
/Note="THE1C-internal repeat: matches 4..1580 of
```

```
consensus"
25314..25670
/Note="MSTA repeat: matches 1..426 of consensus"
repeat_region 25824..26229
/Note="MSTC repeat: matches 1..405 of consensus"
repeat_region 26294..26502
/Note="MER20 repeat: matches 4..218 of consensus"
repeat_region 26762..26813
/Note="26 copies 2 mer tt 78% conserved"
repeat_region 28432..28820
/Note="IIMEC repeat: matches 274..660 of consensus"
repeat_region 28915..29086
/Note="MER30 repeat: matches 4..189 of consensus"
misc_feature 29497..29907
complement(29497..29907)
/Note="match: GSS: Em:AQ022852"
repeat_region 29557..29600
/Note="11 copies 4 mer ttca 81% conserved"
misc_feature 29680..29873
/Note="match: STS: Em:G43807"
repeat_region 30511..30631
/Note="WIR repeat: matches 47..167 of consensus"
repeat_region 30837..31018
/Note="WIR repeat: matches 5..197 of consensus"
misc_feature 32337..32620
complement(32337..32620)
/Note="match: GSS: Em:AQ211265"
repeat_region 32610..32993
/Note="MLT1B repeat: matches 1..390 of consensus"
misc_feature 32626..33078
/Note="match: GSS: Em:AQ268926"
repeat_region 33111..33163
/Note="L2 repeat: matches 2635..2682 of consensus"
repeat_region 33324..33363
/Note="WIR repeat: matches 224..262 of consensus"
repeat_region 33364..33661
/Note="AluSc repeat: matches 3..300 of consensus"
repeat_region 33662..33812
/Note="WIR repeat: matches 67..224 of consensus"
misc_feature 33756..34401
/Note="match: GSS: Em:AQ581639"
repeat_region 34180..34483
/Note="AluSp repeat: matches 3..307 of consensus"
repeat_region 35381..35451
/Note="MER2 repeat: matches 269..343 of consensus"
repeat_region 35481..35697
/Note="AluJb repeat: matches 82..298 of consensus"
repeat_region 35698..35889
/Note="MER2 repeat: matches 16..207 of consensus"
repeat_region 36055..36384
/Note="WIR repeat: matches 56..410 of consensus"
repeat_region 36649..36726
/Note="WIR repeat: matches 65..141 of consensus"
repeat_region 38003..38382
/Note="L2 repeat: matches 2033..2458 of consensus"
repeat_region 38473..38568
```

Query Match 1.4%; Score 47; DB 9; Length 170632;
Best Local Similarity 100.0%; Pred.No.5.2e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3397 AAATGGCAAA 3443
Db 11229 AAATGGCAAA 11275

RESULT 23

LOCUS BV047979
DEFINITION S209P6432FB11.T0 C3H/HeJ Mus musculus STS genomic, sequence tagged site.
ACCESSION BV047979
VERSION BV047979.1 GI:31163774
KEYWORDS STS.
SOURCE Mus musculus (house mouse)

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 467)

REFERENCE
Wade,C.M., Kulbokas,E.J. III, Kirby,A.W., Zody,M.C., Mullikin,J.C.,
Lander,E.S., Lindblad-Toh,K. and Daly,M.J.
The mosaic structure of variation in the laboratory mouse genome
Nature 420 (6915), 574-578 (2002)

JOURNAL
MEDLINE
22354684
PUBMED
12466852

COMMENT
Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersli@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 467
Protocol:
WGS-discovery: Paired-end low-coverage whole genome shotgun reads
were generated from 129S1/SvImJ, C3H/HeJ, and BALB/cByJ. The WGS
reads were placed uniquely on the MGCv3 C57BL/6J assembly and SNP
detection was carried out by SSAHA-SNP. 225,000 reads were
annotated
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J
and the strain from which the particular read came. The validation
rate for these SNPs was estimated at approximately 98%.

FEATURES
source
1..467
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C3H/HeJ"
/db_xref="taxon:10090"
/map="4 15 19-364 56751392-56751733"
/clone_lib="C3H/HeJ"
<1..>467

STTS
ORIGIN

Query Match 1.3%; Score 46; DB 11; Length 467;
Best Local Similarity 100.0%; Pred. No. 2.5e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3398 AATGGCAA 3443
|||||
Db 88 AATGGCAA 133
|||||

RESULT 24
BC043523
LOCUS
DEFINITION
Homo sapiens, clone IMAGE:5166551, mRNA.
ACCESSION
BC043523
VERSION
BC043523.1 GI:27694178
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 682)

AUTHORS
Strausberg,R.
Direct Submission
Submitted (09-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

JOURNAL
TITLE
NTH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Aeano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhur, Parvaneh Saesdi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Scott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 87 Row: f Column: 4.

FEATURES
source
1..682
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5166551"
/issue_type="Brain, adult medulla"
/clone_lib="NIH MGC_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORTc"

ORIGIN

Query Match 1.3%; Score 46; DB 9; Length 682;
Best Local Similarity 100.0%; Pred. No. 2.4e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3398 AATGGCAA 3443
|||||
Db 594 AATGGCAA 639
|||||

RESULT 25
BC053171
LOCUS
DEFINITION
Danio rerio zgc:63964, mRNA (CDNA clone MGC:63964 IMAGE:6791988),
complete cds.
ACCESSION
BC053171
VERSION
BC053171.1 GI:31418813
KEYWORDS
MGC.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE
1 (bases 1 to 1412)

AUTHORS
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heide,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McSwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Maman,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalios,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

TITLE
JOURNAL
PROC. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED
12477932

```

REFERENCE 2 (bases 1 to 1412)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
cdna Library Preparation: Invitrogen Corp
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 117 Row: e Column: 23
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.
FEATURES
source
1. .1412
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MGC:63964 IMAGE:6791988"
/tissue_type="Kidney, zebrafish"
/clone_lib="NCI CGAP_ZK1d1"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6.1"
1. .1412
/gene="zgc:63964"
/notes="synonym: MGC63964"
/db_xref="LocusID:393323"
196..1275
/gene="zgc:63964"
/codon_start=1
/product="hypothetical protein MGC63964"
/protein_id="AAH53171.1"
/db_xref="GI:31418814"
/db_xref="LocusID:393323"
/translacion="MPAWVDKYGKNDVLRFTKNAALPIIHYNEVVVKVHAAGLNPI
DISMRGIGTAATMAKRDPLNISQSGGEFFLLIGRUVSGEIMEGLDVKYFKPGDQVM
AATPPWQSGSLAEFVVVSGNEVSHKPSLRHDEAASIPYVAATAWSAIVNTGGKDN
SAKKRVILGSGGVGTFAIQMKVKAHGVTVTCQNAERLVRDLGADDDVVDYTAGPV
EKOLKNLEKFDLILDSIGGTEKWLALDKPWSGAKFVTLITPLQNTDRGLGADGMW
QSAVTVCVKVKNLKGHVHWGFPAPSGSALDEVSEWVDAGKVRPVVEVFSQAQVP
EAFQKVRQGHARGKTIVSIMEQKE"
ORIGIN
Query Match 1.3%; Score 46; DB 5; Length 1412;
Best Local Similarity 100.0%; Pred. No. 2.3e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3398 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 1363 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1408
RESULT 26
BC017176 1415 bp mRNA linear PRI 29-JUN-2004
DEFINITION Homo sapiens inositol(1myo)-1(or 4)-monophosphatase 2, mRNA (cdna
clone MGC:795 IMAGE:3345789), complete cds.
ACCESSION BC017176
VERSION BC017176.1 GI:16877898
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1415)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Udén, T.B., Toshitoki, S.,
Carninci, P., Prange, C., Raha, S.S., Iqbal, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1415)
Strausberg, R.
Direct Submission
Submitted (05-NOV-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cdna Library Preparation: Rubin Laboratory
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@hgrl.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 5 Row: c Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7657235.
FEATURES
source
1. .1415
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:795 IMAGE:3345789"
/tissue_type="Skin, melanotic melanoma."
/clone_lib="NIH_MGC_20"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
1. .1415
/gene="IMPA2"
/db_xref="LocusID:3613"
/db_xref="MIM:605922"

```

CDS	83..949 /genes="IMPA2" /codon_start=1 /product="inositol(myo)-1(or 4)-monophosphatase 2" /protein_id="AAH17176.1" /db_xref="GI:16877899" /db_xref="LOCUSID:3613" /db_xref="MIN:605922" /translations="MKPSGEDQALAAAGPWEECFQAAVALFAGQIIRKALTEEKRV STKTSADLVTEETHLVEDLIISLERPFPSPHPIAEAAASGAKCVLTHSPWIIDP IDTCNFRHFPFPAVSIIGFAVPOELERGVIVHCTERYLTPRGGRGACNGORLEVS GETDLSKALVLTIGPKRPDPATKFLSNWERLLHAKAGVRVIGSGSTLALCHASGA ADAYYQGLHCWDLDAATAATVIIRAGGIVIDTSGGPLULMACRVVAASTREMAWLIQAQ LQTYNGRDEK"
ORIGIN	Query Match 1.3%; Score 46; DB 9; Length 1415; Best Local Similarity 100.0%; Pred. No. 2.3e-12; Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0; QY 3398 AATGGCAAA 3443 DB 1363 AATGGCAAA 1408
RESULT 27	
BC019066	
LOCUS	BC019066 1705 bp mRNA linear PRI 30-JUN-2004
DEFINITION	Homo sapiens solute carrier family 10 (sodium/bile acid cotransporter family), member 4, mRNA (CDNA clone MGC:29802 IMAGE:4897519), complete cds.
ACCESSION	BC019066
VERSION	BC019066.2 GI:33991607
KEYWORDS	MGC.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	1 (bases 1 to 1705) Klausner,R.D., Collins,F.S., Wagner,L.H., Shemen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Bueto,K.H., Schaefel,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Casavant,T.L., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smalius,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 12477932
TITLE	2 (bases 1 to 1705) Strausberg,R. Direct Submission Submitted (07-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA NIH-MGC Project URL: http://mgc.nci.nih.gov On Aug 20, 2003 this sequence version replaced gi:17512161. Contact: MGC help desk Email: cgaps@mail.nih.gov Tissue Procurement: ATCC
JOURNAL	
PUBLISHED	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
REMARK	
COMMENT	

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E.B. Consortium (LMML)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hghri.nih.gov
Akhter,N., Ayle,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.-L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgenev,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E.B. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 39 Row: j Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 24308413.

FEATURES
source
Location/Qualifiers
1..1705
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="MGC:29802 IMAGE:4897519"
/issue_type="Brain, neuroblastoma"
/clone_lib="NIH MGC_19"
/lab_host="DH10B-R_1"
/note="Vector: pOTB7"
1..1705
/gene="SLC10A4"
/note="synonym: MGC29802"
/db_xref="LOCUSID:201780"
131..1444
/genes="SLC10A4"
/codon_start=1
/product="solute carrier family 10 (sodium/bile acid cotransporter family), member 4"
/protein_id="AAH19066.1"
/db_xref="GI:17512162"
/db_xref="LOCUSID:201780"
/translation="MDGNDNVLLFAPLRDNYTLAPNASSLPGGTDLALAPASSAGP
HGLSLGPGFSPGFPPTPTTSLAGGAASGSPFPFPPHAPHALPFWDTPLN
HGLNVFGAALCTMLGLCTVDVNHFGHVRPVGALLAALCQGLPLPFLALLA
FKLDEAAVAVLLCGCCGGLSNLSLLVDGDMNLSIIMTISLLALVPLCLMW
YSWAWINTPVOLLPGTTLTLCSTLPIGLGVETRYKYSRVADYIVKVSLSLVT
LVLFMTGTMLGPELLASIPAAVYVIALFMPLAGYASGYGLATLPHLPNCRTVCL
ETGSQNVQCTAILKLAFPPQFTGSMFPLLYALFQSAEAGIFVLIYKMYGSEMLHK
RDPLDEDEDTISYKKLKEEADTSYGTVAENIIMETAQTSL"

ORIGIN
Query Match 1.3%; Score 46; DB 9; Length 1705;
Best Local Similarity 100.0%; Pred. No. 2.3e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3398 AATGGCAAA 3443
|||||
DB 1647 AATGGCAAA 1692
|||||

RESULT 28
AK026494
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AK026494 2258 bp mRNA linear PRI 13-SEP-2003
Homo sapiens cDNA: FLJ22841 fis, clone KATA4844.
AK026494
GI:10439367
oligo capping; fis (full insert sequence).
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1
REFERENCE
AUTHORS
Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T.,
Matsumura,K., Nakaajima,Y., Mizuno,T., Morinaga,M., Tanigami,A.,
Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hiraio,M.,
Ohmori,Y., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T.,
Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2258)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail: flicdn@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

TITLE
JOURNAL
AUTHORS
COMMENT

FEATURES
source
Location/Qualifiers
1. .2258
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="KA1A484"
/tissue_type="ileal mucosa"
/clone_lib="kaia"
/note="cloning vector pME18SPL3"

ORIGIN
Query Match 1.3%; Score 46; DB 9; Length 2258;
Best Local Similarity 100.0%; Pred. No. 2.3e-12; Indels 0; Gaps 0;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
3398 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|||||
2201 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2246
|||||

RESULT 29
BC082653 2432 bp mRNA linear VRT 16-SEP-2004
LOCUS
Xenopus laevis cDNA clone MGC:81937 IMAGE:7009628, complete cds.
DEFINITION
ACCESSION BC082653
VERSION BC082653.1 GI:52138921
KEYWORDS MGC.
SOURCE
ORGANISM
Xenopus laevis
Xenopus laevis (African clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 2432)
Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
and Richardson,P.
Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative
Dev. Dyn. 225 (4), 384-391 (2002)
1245917
2 (bases 1 to 2432)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,P.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Ahtshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heide,F.,
Diatchenko,L., Narusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Uedin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,

McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettaman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.C., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalusz,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
3 (bases 1 to 2432)
Klein,S. and Gerhard,D.S.
Direct Submission
Submitted (15-SEP-2004) National Institutes of Health, Xenopus Gene
Collection (XGC), National Institute of Child Health and Human
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
20892-7510, USA
NIH-MGC Project
Contact: XGC help desk
Email: ggapbs-rc@mail.nih.gov
Tissue Procurement: Dr. Igor Dawid
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhbu,
Parvaneh Saeedi, JR Santos, Angeliue Schnerch, Ursula Skalska,
Duane Smalusz, Jeff Scott, Miranda Tsai, George Yang, Jacqui
Schein, Asim Siddiqui, Rob Holt, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 162 Row: c Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.
FEATURES
source
Location/Qualifiers
1. .2432
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="MGC:81937 IMAGE:7009628"
/tissue_type="Kidney, adult Xenopus"
/clone_lib="NICHG_XGC_Kidl"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
51. .1247
/codon_start=1
/product="Unknown (protein for MGC:81937)"
/protein_id="AAH82653.1"
/db_xref="GI:52138922"
/translation="MDGFQGGFKHFAEIHVGOEPEOWKSMVAVPISLSTFTKQFG
PBGHAGDYRSNGNPTNRSNEKVAALDGKAYCLSYASGLAATINTHLKCGDQIC
TKDVGCTNRYFRVASEMGFKVAFVDCSLKLEAATITDTKLWITETPTPLTVI
DIKCGADIVHKKHDIILAVNDTFLSAYFORLALGADI CMYATKYNMGNSDVVMGLV
SVNCDKLYEKLQNSLGAIPSPFPCDKNCRKLTHLRMOHQHFNALTVAKFLEAD
PRVDKLVLPGLSPHQVELTKRQCTGVSGMITFYIKGTLEHAKTFLKSIKVFALAEEL
GGYSLAEHPAINTHASVPEERAALNISTLRSLVGLDEDVEDLIEDLQALGSAHP
DIKNHN"

ORIGIN
Query Match 1.3%; Score 46; DB 5; Length 2432;
Best Local Similarity 100.0%; Pred. No. 2.3e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3398 AATGGCAA 3443
|||||
Db 2363 AATGGCAA 2408

RESULT 30

BC018294

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

BC018294 2500 bp mRNA linear ROD 29-JUN-2004
Mus musculus RIKEN cDNA 1110061004 gene, mRNA (cDNA clone MGC:19290
IMAGE:4021243), complete cds.

BC018294

BC018294.1 GI:17390693

MGC.

Mus musculus (house mouse)

Mus musculus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 2500)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Altschul,S.F., Zordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P.,

Hopkins,R.F., Jordan,H., Farmer,A.A., Rubin,G.M., Hong,L.,

Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,

Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,

Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,

Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,

Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,

McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

Worley,K.C., Hale,K.A., Garcia,A.M., Gay,L.J., Hulyk,S.W.,

Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,

Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,

Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,

Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,

Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalusz,D.E.,

Schmerch,A., Schein,J.E., Jones,S.O. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2500)

Strausberg,R.

Direct Submission

Submitted (03-DEC-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapsb-x@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome

Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegod, H.,

Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,

A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 24 Row: b Column: 11

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis, GenomeScan gene prediction.

Location/Qualifiers

1. .2500

/organism="Mus musculus"

/mol_type="mRNA"

/strain="CZECH II"

gene

CDS

ORIGIN

Query Match

Best Local Similarity

Matches

Score

DB

Length

2500;

Pred. No.

2.3e-12;

Mismatches

0;

Indels

0;

Gaps

0;

QY

3398

AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443

|||||

Db

2431

AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2476

|||||

Search completed: March 12, 2005, 05:53:47

Job time : 9668 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 23:55:25 ; Search time 1165 Seconds
(without alignments)
17494.995 Million cell updates/sec

Title: US-10-627-132-29

Perfect score: 3443

Sequence: 1 ctgcgtgcctgcctgcga.....aaaaaaaaaaaaaaaaaaaa 3443

Scoring table:

OLIGO NUC
Gapop=60, Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 25

Total number of hits satisfying chosen parameters: 105744

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : N_Geneseq_16Dec04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	1.5	1189	3	Aaz58264 Corn cell
2	48	1.4	907	9	Acd44873 Human sig
3	48	1.4	1795	6	Abq54275 Human ova
4	48	1.4	2218	12	Adq24226 Human sof
5	47	1.4	1826	4	Aah33312 Human col
6	45	1.3	189	3	Aaf22014 Human bre
7	45	1.3	480	9	Ach24208 Human adu
8	45	1.3	986	2	Aax51735 DNA encod
9	45	1.3	986	6	Abq92587 Human sec
10	45	1.3	1434	6	Abk33557 cDNA enco
11	45	1.3	1434	8	ACA68518 Novel hum
12	45	1.3	1434	9	ABT44247 Human PRO
13	45	1.3	1434	9	ABT44530 Human PRO
14	45	1.3	1434	9	ACD82197 Human sec
15	45	1.3	1434	9	ABT43903 Human mem
16	45	1.3	1434	9	ADB83533 Novel hum
17	45	1.3	1434	9	ADB80639 Novel hum
18	45	1.3	1434	9	ADB73180 Novel hum
19	45	1.3	1434	9	ADB78262 Novel hum
20	45	1.3	1434	10	ADB84910 Human PRO

21	45	1.3	1434	10	ADB78016	Novel hum
22	45	1.3	1434	10	ADB87082	Human PRO
23	45	1.3	1434	10	ADB84664	Human PRO
24	45	1.3	1434	10	ADB83779	Novel hum
25	45	1.3	1434	10	ADB72934	Novel hum
26	45	1.3	1434	10	ADC36772	Human PRO
27	45	1.3	1434	10	ADC21762	Human PRO
28	45	1.3	1434	10	ADC49793	Novel hum
29	45	1.3	1434	10	ADC48992	Novel hum
30	45	1.3	1434	10	ADC49509	Novel hum
31	45	1.3	1434	10	ADC47370	Novel hum
32	45	1.3	1434	10	ADC47115	Novel hum
33	45	1.3	1434	10	ADC77990	Novel hum
34	45	1.3	1434	10	ADD06225	Novel hum
35	45	1.3	1434	10	ADC77744	Novel hum
36	45	1.3	1434	10	ADD50707	Novel hum
37	45	1.3	1434	10	ADD50953	Novel hum
38	45	1.3	1434	10	ADD50434	Human PRO
39	45	1.3	1434	10	ADD50188	Human PRO
40	45	1.3	1434	10	ADD51199	Novel hum
41	45	1.3	1434	10	ADB89052	Encoding
42	45	1.3	1434	10	ACA66862	Acad66862 cDNA enco
43	45	1.3	1434	10	ACD68614	Novel hum
44	45	1.3	1434	12	ADC48746	Novel hum
45	45	1.3	1434	12	ADE20917	Novel hum
46	45	1.3	1434	12	ADE05761	Human PRO
47	45	1.3	1434	12	ADD74990	Human PRO
48	45	1.3	1434	12	ADD75736	Novel hum
49	45	1.3	1434	12	ADD84968	Novel hum
50	45	1.3	1434	12	ADD86794	Novel hum
51	45	1.3	1434	12	ADE20671	Novel hum
52	45	1.3	1434	12	ADE38968	Novel hum
53	45	1.3	1434	12	ADE05515	Human PRO
54	45	1.3	1434	12	ADD73500	Human PRO
55	45	1.3	1434	12	ADD78340	Novel hum
56	45	1.3	1434	12	ADE21163	Novel hum
57	45	1.3	1434	12	ADD77278	Novel hum
58	45	1.3	1434	12	ADE20425	Novel hum
59	45	1.3	1434	12	ADD75490	Human PRO
60	45	1.3	1434	12	ADD74006	Human PRO
61	45	1.3	1434	12	ADD74252	Human PRO
62	45	1.3	1434	12	ADD75982	Novel hum
63	45	1.3	1434	12	ADD85474	Novel hum
64	45	1.3	1434	12	ADE05023	Human PRO
65	45	1.3	1434	12	ADD75236	Human PRO
66	45	1.3	1434	12	ADD76780	Novel hum
67	45	1.3	1434	12	ADD86548	Novel hum
68	45	1.3	1434	12	ADD78016	Novel hum
69	45	1.3	1434	12	ADD77524	Novel hum
70	45	1.3	1434	12	ADD77770	Novel hum
71	45	1.3	1434	12	ADD85228	Novel hum
72	45	1.3	1434	12	ADD73760	Human PRO
73	45	1.3	1434	12	ADD74498	Human PRO
74	45	1.3	1434	12	ADD77026	Novel hum
75	45	1.3	1434	12	ADD85720	Novel hum
76	45	1.3	1434	12	ADE05269	Human PRO
77	45	1.3	1434	12	ADD74744	Human PRO
78	45	1.3	1434	12	ADG05556	Novel hum
79	45	1.3	1434	12	ADG27110	Human PRO
80	45	1.3	1434	12	ADG11173	Novel hum
81	45	1.3	1434	12	ADG11952	Novel hum
82	45	1.3	1434	12	ADG94509	Novel hum
83	45	1.3	1434	12	ADG06605	Human PRO
84	45	1.3	1434	12	ADH38949	Novel hum
85	45	1.3	1434	12	ADG34039	Novel hum
86	45	1.3	1434	12	ADH33509	Human PRO
87	45	1.3	1434	12	ADH69603	Human PRO
88	45	1.3	1434	12	ADI29764	Novel hum
89	45	1.3	1434	12	ADM27161	Novel hum
90	45	1.3	1434	12	ADK66519	Human PRO
91	45	1.3	1576	5	AAD07711	Human sec
92	45	1.3	1772	3	AAF21809	Human nuc
93	45	1.3	1947	8	AAD55845	Human nuc

c 94 45 1.3 1984 3 AAC59248 Aac59248 Human sec
95 45 1.3 2592 12 ADQ22667 Adq22667 Human sof
96 45 1.3 3118 3 AAF15989 Aaf15989 Human pro
97 45 1.3 3145 3 AAC59241 Aac59241 Human sec
98 45 1.3 3872 10 ADE79063 Ade79063 Human pro
99 45 1.3 4125 13 ADR44022 Adr44022 Human col
c 100 44 1.3 161 6 ABK54773 Abk54773 Human col

ALIGNMENTS

RESULT 1
AAZ58264
ID AAZ58264 standard; cDNA; 1189 BP.
XX AC AAZ58264;
XX AC AAZ58264;
DT 08-MAY-2000 (first entry)
XX Corn cellulose synthase cDNA crin.pk0135.e10.
XX Corn; maize; cellulose synthase; transgenic plant; ss.
XX Zea mays.
XX Key Location/Qualifiers
FH 2. .964
FT CDS /*tag= a
FT /partial
XX WO200004166-A2.
XX 27-JAN-2000.
XX 13-JUL-1999; 99WO-US015871.
XX 14-JUL-1998; 98US-0092844P.
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G;
PI Rafaleki JA, Thorpe CJ;
XX WPI; 2000-182431/16.
XX P-PSDB; AAY58833.
XX Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as
PT probes for isolating cDNAs and genes encoding homologous proteins, for
PT producing transgenic plants.

XX Claim 3; Page 42-43; 93pp; English.
XX The present sequence is that of cDNA clone crin.pk0135.e10 encoding a
CC portion (see AAY58833) of corn cellulose synthase (CS). The cDNA clone
CC was isolated from a seedling root cDNA library on the basis of homology
CC to Arabidopsis and cotton CS sequences. The invention relates to isolated
CC nucleic acid fragments encoding plant CS and to CS polypeptides. It also
CC relates to the construction of a chimeric gene encoding all or a portion
CC of the CS, in sense or antisense orientation, where expression of the
CC gene results in altered levels of the CS in transformed host cells. The
CC host cells can be used to screen compounds for their ability to inhibit
CC CS activity. CS nucleic acids are also useful for producing transgenic
CC plants having altered levels of CS, and hence altered levels of fibre. CS
CC may also serve as a target for the development of novel herbicides
XX
SQ Sequence 1189 BP; 215 A; 374 C; 331 G; 269 T; 0 U; 0 Other;
Query Match 1.5%; Score 50; DB 3; Length 1189;
Best Local Similarity 100.0%; Pred. No. 3.9e-08;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2771 AACGAGCAGTTCTGGGTGCATCGCGCGGTGTCGCGCATCTCTTCGCCGT 2820

Db 485 AACGAGCAGTTCTGGGTGCATCGCGCGGTGTCGCGCATCTCTTCGCCGT 534
RESULT 2
ACD44873
ID ACD44873 standard; cDNA; 907 BP.
XX AC ACD44873;
XX AC ACD44873;
DT 10-SEP-2003 (first entry)
XX Human signal transduction pathway component HDPSE86 cDNA.
XX Human; signal transduction pathway component; gene; ss; gene therapy;
KW neural disorder; immune system disorder; systemic lupus erythematosus;
KW rheumatoid arthritis; multiple sclerosis; muscular disorder; sinusitis;
KW respiratory disease; nasal vestibulitis; nasal polyp; pulmonary disorder;
KW cardiovascular disorder; congenital heart defect; Ebstein's anomaly;
KW hypoplastic left heart syndrome; renal disorder; acute kidney failure;
KW end-stage renal disease; hyperproliferative disorder; Hodgkin's disease;
KW leukaemia; inflammatory disease; septic shock; bursitis; appendicitis;
KW allergy; asthma; blood related disorder; thrombosis; atherosclerosis;
KW myocardial infarction; endocrine disorder; Addison's disease; dysphagia;
KW corticosteroid deficiency; reproductive system disorder; dysmenorrhea;
KW testicular atrophy; gastrointestinal disorder; irritable bowel syndrome;
KW infectious disease; cancer; wound healing; epithelial cell proliferation.
XX Homo sapiens.
OS US2003036505-A1.
PN 20-FEB-2003.
XX 20-SEP-2001; 2001US-00955999.
XX 25-SEP-2000; 2000US-0234997P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Barash SC, Ni J, Ruben SM, Rosen CA, Young PE, Rohrschneider LR;
PI WPI; 2003-492122/46.
XX P-PSDB; ABO27187.
XX New isolated nucleic acids encoding signal transduction pathway component
PT polypeptides, useful for diagnosing, treating, and/or preventing
PT disorders, such as cancer, infections, cardiovascular and inflammatory
PT diseases.
XX Claim 3; Page 211; 297pp; English.
XX The invention relates to an isolated nucleic acid molecule. The methods
CC and compositions of the present invention are useful for diagnosing,
CC treating, preventing and/or prognosing disorders related to the novel
CC polypeptides, such as neural disorders, immune system disorders (e.g.
CC systemic lupus erythematosus, rheumatoid arthritis, or multiple
CC sclerosis), muscular disorders, respiratory diseases (e.g. nasal
CC vestibulitis, nasal polyps, or sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly or hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure, or end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease or
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis, or
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis or myocardial
CC infarction), endocrine disorders (e.g. Addison's disease or
CC corticosteroid deficiency), reproductive system disorders (e.g.
CC testicular atrophy or dysmenorrhea), gastrointestinal disorders (e.g.
CC dysphagia or irritable bowel syndrome), infectious diseases, and/or
CC cancerous diseases. The polynucleotides can also be used to aid wound
CC healing and epithelial cell proliferation. The present sequence
CC represents a human signal transduction pathway component cDNA

XX Sequence 907 BP; 211 A; 264 C; 255 G; 173 T; 0 U; 4 Other;
SQ
Query Match 1.4%; Score 48; DB 9; Length 907;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3396 TAAATGCGCAAA 3443
Db 859 TAAATGCGCAAA 906
RESULT 3
ABQ54275
ID ABQ54275 standard; cDNA; 1795 BP.
XX AC ABQ54275;
XX XX
XX 22-AUG-2002 (first entry)
XX Human ovarian antigen HDPSE86 cDNA, SEQ ID NO:155.
XX
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive; chromosome 12q23-24;
KW gene; ss.
XX
XX Homo sapiens.
OS
XX
XX W0200200677-A1.
PN
XX
XX 03-JAN-2002.
PD
XX
XX 07-JUN-2001; 2001WO-US018569.
XX
XX 07-JUN-2000; 2000US-0209467P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Birse CE, Rosen CA;
PI
XX
XX WPI; 2002-147878/19.
DR
XX P-PSDB; ABP41198.
DR
XX
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
PT cancer), immune disorders, cardiovascular disorders and neurological
PT diseases.
XX
XX Claim 1; SEQ ID NO 155; 2922bp; English.
PS
XX
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired

CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders,
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents cDNA encoding a human ovarian antigen of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1795 BP; 370 A; 548 C; 530 G; 341 T; 0 U; 6 Other;
Query Match 1.4%; Score 48; DB 6; Length 1795;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3396 TAAATGCGCAAA 3443
Db 1747 TAAATGCGCAAA 1794
RESULT 4
ADQ24226
ID ADQ24226 standard; DNA; 2218 BP.
XX AC ADQ24226;
XX
XX 26-AUG-2004 (first entry)
XX
XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 7046.
XX
XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.
XX
XX Homo sapiens.
OS
XX
XX W02004048938-A2.
PN
XX
XX 10-JUN-2004.
PD
XX
XX 26-NOV-2003; 2003WO-US038193.
XX
XX 26-NOV-2002; 2002US-0429739P.
XX
XX (PROT-) PROTEIN DESIGN LABS INC.
XX
XX Aziz N, Ginsburg WM, Zlotnik A;
PI
XX WPI; 2004-441208/41.
DR
XX
XX Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
XX Example 2; SEQ ID NO 7046; 210pp; English.
PS
XX
XX The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual,
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated

CC DNA of the invention. The current sequence is not shown within the
 CC specification per se but was submitted in CD format by the inventor.

XX Sequence 2218 BP; 499 A; 636 C; 646 G; 437 T; 0 U; 0 Other;

SQ Query Match 1.4%; Score 48; DB 12; Length 2218;

Best Local Similarity 100.0%; Pred. No. 1.7e-07;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3396 TAAATGGCAA 3443

DB 1754 TAAATGGCAA 1801

RESULT 5

AAH33312

ID AAH33312 standard; cDNA; 1826 BP.

XX

AC AAH33312;

XX 03-SEP-2001 (first entry)

DE Human colon cancer antigen encoding cDNA SEQ ID NO:368.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;

KW colorectal carcinoma; chromosome 12; ss.

XX Homo sapiens.

OS

XX WO200122920-A2.

PN

XX 05-APR-2001.

PD

XX 28-SEP-2000; 2000NO-US026524.

PF

XX 29-SEP-1999; 99US-0157137P.

PR

XX 03-NOV-1999; 99US-0163280P.

PR

XX (HUMA-) HUMAN GENOME SCI INC.

XX

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

PI WPI; 2001-235357/24.

DR P-PSDB; AAG73881.

DR

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,

XX useful for preventing, diagnosing and/or treating colorectal cancers.

XX

XX Claim 1; Page 2476-2477; 9803pp; English.

XX

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon

XX cancer-associated nucleic acid molecules (N) and proteins (P), where the

XX proteins are collectively known as colon cancer antigens. The colon

XX cancer antigens have cytostatic activity and can be used in gene therapy

XX and vaccine production. N and P may be used in the prevention, diagnosis

XX and treatment of diseases associated with inappropriate P expression. For

XX example, N and P may be used to treat disorders associated with decreased

XX expression by rectifying mutations or deletions in a patient's genome

XX that affect the activity of P by expressing inactive proteins or to

XX supplement the patient's own production of P. Additionally, N may be used

XX to produce the colon cancer-associated Ps, by inserting the nucleic acids

XX into a host cell and culturing the cell to express the proteins. N and P

XX can be used in the prevention, diagnosis and treatment of colorectal

XX carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent

XX sequences used in the exemplification of the present invention. N.B.

XX Pages 666 to 682 and page 7053 of the sequence listing were missing at

XX time of publication, meaning no sequences are present for SEQ ID NO:1027

XX to 1052, 7921 and 7922

XX

SQ Sequence 1826 BP; 593 A; 360 C; 318 G; 554 T; 0 U; 1 Other;

Query Match

Best Local Similarity 1.4%; Score 47; DB 4; Length 1826;

Pred. No. 4.1e-07;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3397 AAATGGCAA 3443

DB 1740 AAATGGCAA 1786

RESULT 6

AAF22014

ID AAF22014 standard; DNA; 189 BP.

XX

AC AAF22014;

XX

DT 27-MAR-2001 (first entry)

XX

XX Human breast and ovarian cancer associated antigen gene SEQ ID 401.

XX

KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;

KW nontropic; neuroprotective; antiviral; antiallergic; hepatotropic;

KW antidiabetic; antiinflammatory; antiulcer; vulnery; anticonvulsant;

KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;

KW Addison's disease; allergy; autoimmune haemolytic anaemia;

KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;

KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;

KW cardiovascular disorder; wound healing; neurological disease; ds.

XX

XX Homo sapiens.

XX

XX WO200055173-A1.

PN

XX 21-SEP-2000.

PD

XX 08-MAR-2000; 2000WO-US005881.

PF

XX 12-MAR-1999; 99US-0124270P.

PR

XX (HUMA-) HUMAN GENOME SCI INC.

XX

XX Rosen CA, Ruben SM;

XX

XX WPI; 2000-511515/58.

DR P-PSDB; AAB59111.

DR

XX New human breast and ovarian cancer associated gene sequences and the

XX polypeptides encoded by these genes, useful in the prevention, treatment

XX and diagnosis of cancer, immune disorders, cardiovascular disorders and

XX neurological diseases.

XX

XX Claim 1; Page 833; 1299pp; English.

XX

XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human

XX proteins AAB58711 - AAB59128. The DNA and protein sequences are

XX associated with breast and ovarian cancer. Included in the invention are

XX sequences AAF22032 - AAF22040 and AAB59129 which are used in the

XX isolation and characterisation of the DNA and protein sequences of the

XX invention. The breast and ovarian cancer associated DNA, protein, agonist

XX or antagonist sequences exhibit cytostatic; immunosuppressive; nontropic;

XX neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;

XX antiinflammatory; antiulcer; vulnery; anticonvulsant; antibacterial;

XX antifungal; antiparasitic and cardiant activity. The polynucleotide and

XX protein sequences are used in the diagnosis of cancer, particularly

XX breast and ovarian cancer. The nucleic acid sequences, proteins, agonists

XX and agonists may also be used in the diagnosis, prevention and treatment

XX of immune disorders e.g. Addison's disease, allergies, autoimmune

XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus. Crohn's

XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;

XX cardiovascular disorders such as myocardial ischaemias; wound healing;

XX neurological diseases such as cerebral anoxia and epilepsy; and

XX infectious diseases

XX

SQ Sequence 189 BP; 87 A; 36 C; 20 G; 40 T; 0 U; 6 Other;

Query Match

1.3%; Score 45; DB 3; Length 189;

Best Local Similarity 100.0%; Pred. No. 3.3e-06; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGCGCAA 3443
104 ATGCGCAA 148

Db

RESULT 7
ACH24208
ID ACH24208 standard; cDNA; 480 BP.
XX
AC ACH24208;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human adult ovary cDNA #2588.
XX
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX
PN US2003073623-A1.
XX
PD 17-APR-2003.
XX
PF 30-JUL-2001; 2001US-00918995.
XX
PR 30-JUL-2001; 2001US-00918995.
XX
PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 11420; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH2789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030073623
XX
SQ Sequence 480 BP; 129 A; 118 C; 118 G; 98 T; 0 U; 17 Other;
Query Match 1.3%; Score 45; DB 9; Length 480;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGCGCAA 3443

Db 208 ATGCGCAA 252

RESULT 8
AAX51735
ID AAX51735 standard; DNA; 986 BP.
XX
AC AAX51735;
XX
DT 17-JUN-1999 (first entry)
XX
DE DNA encoding a human secreted protein.
XX
KW Human secreted protein; cancer; immune disorder; infection;
KW inflammatory disorder; skin disorder; tumour; atherosclerosis;
KW restenosis; autoimmune disorder; Alzheimer's disease;
KW peripheral neuropathy; trauma; spinal cord injury; allergy;
KW hematopoietic disorder; skeletal disorder; neurological disorder;
KW arthritic disorder; asthma; immunodeficiency disease; AIDS;
KW transplant rejection; ss.
XX
OS Homo sapiens.
XX
PN WO9911293-A1.
XX
PD 11-MAR-1999.
XX
PF 03-SEP-1998; 98WO-US018360.
XX
PR 05-SEP-1997; 97US-0057626P.
PR 05-SEP-1997; 97US-0057663P.
PR 05-SEP-1997; 97US-0057669P.
PR 12-SEP-1997; 97US-0058666P.
PR 12-SEP-1997; 97US-0058667P.
PR 12-SEP-1997; 97US-0058973P.
PR 12-SEP-1997; 97US-0058974P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Moore PA, Ruben SM, Lafleur DW, Shi Y, Rosen GA, Olsen HS;
XX Ebner R, Brewer LA;
XX WPI; 1999-204988/17.
XX P-PSDB; AAY12948.
XX
XX New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. neurological disorders,
PT tumors, immune disorders, inflammation or hematological disorders.
XX
XX Claim 1; Page 171; 215pp; English.
XX
XX AAX51701-55 encode human secreted proteins. The polynucleotides and their
CC corresponding secreted polypeptides are useful for preventing, treating
CC or ameliorating medical conditions, e.g. by protein or gene therapy.
CC Pathological conditions can also be diagnosed by determining the amount
CC of the new polypeptides in a sample or by determining the presence of
CC mutations in the new polynucleotides. Specific uses are described for
CC each polynucleotide, based on which tissues they are most highly
CC expressed in, and include developing products for the diagnosis or
CC treatment of cancer, immune disorders, infection, inflammatory disorders,
CC skin disorders, tumors, atherosclerosis, restenosis, autoimmune
CC disorders, Alzheimer's disease, peripheral neuropathies, trauma, spinal
CC cord injuries, allergy, hematopoietic disorders, skeletal disorders,
CC neurological disorders, arthritic disorders, asthma, immunodeficiency
CC diseases, AIDS and transplant rejection. The polypeptides are also useful
CC for identifying their binding partners
XX
SQ Sequence 986 BP; 371 A; 132 C; 177 G; 306 T; 0 U; 0 Other;
Query Match 1.3%; Score 45; DB 2; Length 986;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 3399 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 883 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 9
ABQ92587
ID ABQ92587 standard; cDNA; 986 BP.
XX
AC ABQ92587;
XX
DT 12-NOV-2002 (first entry)
XX
DE Human secreted protein encoding cDNA SEQ ID NO 45.
XX
KW Human; neotropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antiskilling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN WO200257420-A2.
XX
PD 25-JUL-2002.
XX
PF 17-JAN-2002; 2002WO-US001109.
XX
PR 18-JAN-2001; 2001US-0262066P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Moore PA, Ruben SM, Lafleur DW, Shi Y, Rosen CA, Olsen H;
PI Ebner R, Brewer LA;
XX
WPI: 2002-599716/64.
DR P-PSDB; ABP62047.
XX
New polynucleotides and polypeptides useful for diagnosing, prognosing,
PT treating or preventing e.g. neurodegenerative, central nervous system,
PT autoimmune, respiratory, reproductive, or inflammatory diseases or
PT disorders.
XX
PS Claim 1; Page 720; 785pp; English.
XX
CC The invention relates to novel genes (ABQ92553-ABQ92607) and proteins
CC (ABP62013-ABP62153) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
XX infections
XX
SQ Sequence 986 BP; 371 A; 132 C; 177 G; 306 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 6; Length 986;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
```

```
Db 883 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 10
ABK33557
ID ABK33557 standard; cDNA; 1434 BP.
XX
AC ABK33557;
XX
DT 08-MAY-2002 (first entry)
XX
DE cDNA encoding human PRO protein, Seq ID No 43.
XX
KW Human; secreted protein; PRO; tumour; lung cancer; colon cancer;
KW breast cancer; prostate tumour; rectal tumour; liver tumour;
KW pericyte cell proliferation; chondrocyte cell proliferation;
KW tumour necrosis factor-alpha; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200208288-A2.
XX
PD 31-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-US021066.
XX
PR 20-JUL-2000; 2000US-0219556P.
PR 25-JUL-2000; 2000US-0220585P.
PR 25-JUL-2000; 2000US-0220605P.
PR 25-JUL-2000; 2000US-0220607P.
PR 25-JUL-2000; 2000US-0220624P.
PR 25-JUL-2000; 2000US-0220638P.
PR 25-JUL-2000; 2000US-0220664P.
PR 25-JUL-2000; 2000US-0220666P.
PR 26-JUL-2000; 2000US-0220893P.
PR 28-JUL-2000; 2000WO-US020710.
PR 01-AUG-2000; 2000US-0222425P.
PR 22-AUG-2000; 2000US-0227133P.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 10-NOV-2000; 2000WO-US030873.
PR 28-NOV-2000; 2000US-0253646P.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 22-MAR-2001; 2001US-00816744.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001WO-US017092.
XX
PA (GETH ) GENENTECH INC.
XX
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX
WPI: 2002-172001/22.
DR P-PSDB; AAU83613.
XX
One hundred and twenty two nucleic acids encoding PRO polypeptides,
PT useful for treating a PRO related disorder and for diagnosing tumors such
PT as lung cancer, colon cancer, breast tumor, prostate tumor, rectal tumor
PT or liver tumor.
XX
PS Claim 2; Fig 43; 359pp; English.
XX
CC The invention relates to one hundred and twenty two nucleic acids
CC encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides
CC encode human secreted proteins. The PRO nucleic acids, polypeptides,
CC agonists and antagonists are useful for treating a PRO related disorder.
CC The PRO polypeptides are useful for diagnosing tumours, especially lung
```


CC cancer, colon cancer, breast tumour, prostate tumour, rectal tumour or
 CC liver tumour. The PRO polypeptides are useful for stimulating the
 CC proliferation of, or gene expression, in pericyte cells, for stimulating
 CC the proliferation or differentiation of chondrocyte cells, for
 CC stimulating the release of tumour necrosis factor-alpha from human blood,
 CC for stimulating or inhibiting the proliferation of normal human dermal
 CC fibroblast cells. The PRO polypeptide may also be used as molecular
 CC weight markers and for tissue typing. The PRO nucleic acids have
 CC applications in molecular biology, including use as hybridisation probes,
 CC and in chromosome and gene mapping. ABK33536-ABK33657 represent human PRO
 CC protein coding sequences of the invention

SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 6; Length 1434;
 Best Local Similarity 100.0%; Pred. No. 2.2e-06;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAA 3443
 |||||
 Db 1347 ATGGCAA 1391

RESULT 11

ACA68518
 ID ACA68518 standard; cDNA; 1434 BP.

AC ACA68518;

DT 25-JUN-2003 (first entry)

DE Novel human secreted and transmembrane protein PRO9905 cDNA.

KW Human; secreted and transmembrane protein; PRO; cardiant; cytostatic;
 KW antiangiogenic; hypotensive; vulnery; antiarteriosclerotic;
 KW gene therapy; cardiovascular disorder; endothelial disorder;
 KW angiogenic disorder; cardiac hypertrophy; trauma; cancer;
 KW age-related macular degeneration; atherosclerosis; hypertension;
 KW arterial restenosis; rheumatoid arthritis; angina; myocardial infarction;
 KW thrombophlebitis; lymphangitis; tumour angiogenesis; breast carcinoma;
 KW liver carcinoma; wound healing; chromosome mapping; gene mapping; gene;
 KW ss.

OS Homo sapiens.

PN US2003088063-A1.

XX 08-MAY-2003.

PF 12-AUG-2002; 2002US-00219003.

PR 25-JUL-2000; 2000US-0220664P.

PR 01-JUN-2001; 2001WO-US017800.

PR 29-JUN-2001; 2001WO-US021066.

PR 09-APR-2002; 2002US-00119480.

XX (GETH) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;

PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WT;

XX WPI; 2003-393229/37.

DR P-PSDB; ABU82069.

XX One hundred and eighty seven nucleic acids encoding PRO polypeptides,
 XX useful in diagnosis and treatment of cardiovascular (e.g. myocardial
 XX infarction), endothelial or angiogenic disorders in a mammal.

XX Claim 2; Fig 43; 314pp; English.

XX The invention describes one hundred and eighty seven nucleic acids
 CC encoding novel human secreted and transmembrane (PRO) polypeptides. The
 CC PRO nucleic acids, polypeptides, agonists and antagonists are useful for

CC treating or diagnosing a cardiovascular, endothelial or angiogenic
 CC disorder in a mammal, e.g. cardiac hypertrophy, trauma, cancer, age-
 CC related macular degeneration, atherosclerosis, hypertension, arterial
 CC restenosis, rheumatoid arthritis, angina, myocardial infarctions,
 CC thrombophlebitis, lymphangitis, tumour angiogenesis (such as breast
 CC carcinoma and liver carcinoma) and wound healing. The PRO nucleic acids
 CC have applications in molecular biology, including use as hybridisation
 CC probes, and in chromosome and gene mapping. This sequence encodes a novel
 CC human secreted and transmembrane PRO polypeptide

SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 8; Length 1434;
 Best Local Similarity 100.0%; Pred. No. 2.2e-06;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAA 3443
 |||||
 Db 1347 ATGGCAA 1391

RESULT 12

ABT44247

ID ABT44247 standard; cDNA; 1434 BP.

XX ABT44247;

DT 06-NOV-2003 (first entry)

DE Human PRO9905 cDNA.

XX PRO; proliferation; pericyte cell; TNF-alpha; blood; chondrocyte; ss;
 KW differentiation; dermal fibroblast; tumour; gene therapy; gene;
 KW cytostatic.

OS Homo sapiens.

PN US2003050448-A1.

XX 13-MAR-2003.

PF 28-AUG-2002; 2002US-00230414.

PR 01-JUN-2001; 2001WO-US017800.

PR 29-JUN-2001; 2001WO-US021066.

PR 09-APR-2002; 2002US-00119480.

XX (GETH) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;

PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WT;

XX WPI; 2003-521818/49.

DR P-PSDB; ABJ72249.

XX New nucleic acid encoding for a PRO protein, useful for the manufacture
 XX of a medicament for diagnosing or treating tumors or for measuring or
 XX detecting expression of an associated gene.

XX Claim 2; Fig 43; 315pp; English.

XX The invention relates to a novel isolated nucleic acid encoding a fully
 CC defined PRO polypeptide. The molecules of the invention may be useful for
 CC stimulating proliferation or gene expression in pericyte cells or the
 CC release of TNF-alpha from human blood. Other possible uses include the
 CC stimulation or inhibition of chondrocyte proliferation or
 CC differentiation, the stimulation of human dermal fibroblast cell
 CC proliferation and the detection of the presence of a tumour within a
 CC mammal. Furthermore, the nucleic acid may be useful for the manufacture
 CC of a medicament for diagnosing or treating a tumour within a mammal or
 CC for measuring or detecting the expression of an associated gene, as well
 CC as during gene therapy. The current sequence is that of the human PRO
 CC cDNA of the invention

```
XX
SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;
Query Match 1.3%; Score 45; DB 9; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|
Db 1347 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1391

RESULT 13
ABT44530
ID ABT44530 standard; cDNA; 1434 BP.
XX
AC ABT44530;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human PRO9905 cDNA.
XX
KW PRO; proliferation; gene; pericyte cell; TNF alpha; chondrocyte; blood;
KW tumour necrosis factor; proliferation; differentiation; gene therapy;
KW dermal fibroblast; ss.
XX
OS Homo sapiens.
XX
PN US2003027988-A1.
XX
PD 06-FEB-2003.
XX
PF 26-AUG-2002; 2002US-00227884.
XX
PR 01-JUN-2001; 2001WO-US017800.
PR 23-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX
PA (GETH ) GENENTECH INC.
XX
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX
WPI; 2003-503301/47.
P-PSDB; ABJ72377.
XX
PT New PRO protein encoding nucleic acid, useful for preparing PRO
PT polypeptides and anti-PRO antibodies for detecting the presence of a
PT tumor in a mammal.
XX
PS Claim 2; Fig 43; 324pp; English.
XX
CC The invention relates to a novel isolated PRO protein encoding nucleic
CC acid. The nucleic acid of the invention may be useful for preparing PRO
CC polypeptides and anti-PRO antibodies for detecting the presence of a
CC tumour in a mammal. Furthermore, the molecules of the invention may be
CC useful for stimulating proliferation or gene expression in pericyte
CC cells, the release of tumour necrosis factor (TNF)-alpha from human
CC blood, the proliferation or differentiation of chondrocyte cells and for
CC inhibiting the proliferation of normal human dermal fibroblast cells.
CC Finally, the molecules may be utilised during gene therapy. The current
CC sequence is that of the human PRO cDNA of the invention
XX
SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;
Query Match 1.3%; Score 45; DB 9; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|
Db 1347 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1391
```

```
RESULT 14
ACD82197
ID ACD82197 standard; cDNA; 1434 BP.
XX
AC ACD82197;
XX
DT 19-SEP-2003 (first entry)
XX
DE Human secreted/transmembrane polypeptide PRO 9905 cDNA.
XX
KW Human; ss; chondrocyte stimulation; TNF-alpha stimulation; gene therapy;
KW human dermal fibroblast stimulation; tumour; tissue typing; gene;
KW affinity purification.
XX
OS Homo sapiens.
XX
PN US2003044934-A1.
XX
PD 06-MAR-2003.
XX
PF 28-AUG-2002; 2002US-00230338.
XX
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX
PA (GETH ) GENENTECH INC.
XX
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX
WPI; 2003-492274/46.
P-PSDB; ABO34272.
XX
PT New transmembrane polypeptides and nucleic acids encoding the
PT polypeptides, useful in gene therapy, in chromosome identification, as
PT chromosome markers, or in generating probes.
XX
PS Claim 2; Fig 43; 315pp; English.
XX
CC The invention relates to an isolated nucleic acid encoding a PRO
CC polypeptide. Nucleic acids that encode PRO can be used to generate either
CC transgenic animals or knock-out animals useful in developing and
CC screening of therapeutically useful reagents. The nucleic acids may also
CC be used in gene therapy for replacing defective gene, in chromosome
CC identification, as chromosome markers, or in generating probes to isolate
CC full length PRO cDNA. The PRO polypeptides are useful for chondrocyte
CC stimulation, TNF-alpha stimulation, human dermal fibroblasts stimulation
CC and for detecting the presence of tumour in an mammal. The PRO
CC polypeptides are useful as molecular markers for protein electrophoresis
CC and the isolated nucleic acids may be used for recombinantly expressing
CC those markers. The PRO polypeptides and nucleic acids may also be used in
CC tissue typing. Anti-PRO antibodies are useful in diagnostic assays for
CC PRO and in affinity purification of PRO from recombinant cell culture or
CC natural sources. The present sequence represents cDNA encoding a human
CC secreted/transmembrane PRO polypeptide
XX
SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;
Query Match 1.3%; Score 45; DB 9; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|
Db 1347 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1391

RESULT 15
ABT43903
ID ABT43903 standard; cDNA; 1434 BP.
XX
```

AC ABT43903;
XX
DT 16-OCT-2003 (first entry)
XX
DE Human membrane bound receptor/protein PRO9905 cDNA sequence.
XX
KW Human; PRO; membrane bound protein; membrane bound receptor;
KW cell proliferation; cell migration; cell differentiation;
KW mitogenic factor; survival factor; cytotoxic factor;
KW differentiation factor; neuro peptide; hormone; cell receptor;
KW receptor-ligand interaction; cytostatic; chondrocyte; tumour; gene; ss.
XX
OS Homo sapiens.
XX
FN US2003065147-A1.
XX
PD 03-APR-2003.
XX
PF 29-AUG-2002; 2002US-00232224.
XX
PR 28-JUL-1999; 99US-0146222P.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX
PA (GETH) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
PI
XX WPI; 2003-522018/49.
DR P-PSDB; ABJ72079.
XX
XX One hundred and twenty two nucleic acids encoding PRO polypeptides
PT useful for the manufacture of a medicament for diagnosing or treating
PT tumor.
XX
XX Claim 2; Fig 43; 315pp; English.
XX
XX This invention relates to one hundred and twenty two novel nucleic acids
CC encoding human PRO membrane bound proteins or receptors. Extracellular
CC proteins play important roles in the formation, differentiation and
CC maintenance of multicellular organisms. The fate of many individual cells
CC (for example proliferation, migration or differentiation) is typically
CC governed by information received from other cells and the immediate
CC environment. The information is often transmitted by secreted
CC polypeptides (for example mitogenic factors, survival factors, cytotoxic
CC factors, differentiation factors, neuro peptides and hormones) which are
CC received and interpreted by diverse cell receptors or membrane bound
CC proteins. These membrane bound proteins and receptors may be of use as
CC pharmaceutical and diagnostic agents, such as in the blocking of receptor
CC -ligand interactions. The current invention provides the amino acid
CC sequences of novel human membrane bound receptors and proteins, along
CC with the cDNA sequences encoding them. The novel proteins of the
CC invention may have cytostatic activities through the stimulation of
CC chondrocytes. The nucleic acids of the invention may be useful for the
CC manufacture of a medicament for diagnosing or treating a tumour in a
CC mammal. In addition, they may be useful for measuring or detecting the
CC expression of a tumour associated gene. The present sequence is the cDNA
XX sequence encoding a human PRO protein of the invention
SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;
Query Match 1.3%; Score 45; DB 9; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 3399 ATGCGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|||||
Db 1347 ATGCGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1391
|||||

RESULT 16
ADB83533
ID ADB83533 standard; cDNA; 1434 BP.
XX
AC ADB83533;
XX
DT 04-DEC-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO9905 cDNA.
XX
KW human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
KW vulnary; antiarthritic; pericyte cell proliferation;
KW pericyte cell differentiation; chondrocyte cell proliferation;
KW chondrocyte cell differentiation; tumour necrosis factor alpha release;
KW (TNF)-alpha release; dermal fibroblast cell proliferation;
KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW liver tumour; tissue typing; chromosome mapping; gene mapping;
KW gene therapy.
XX
XX Homo sapiens.
XX
XX US2003073814-A1.
XX
PD 17-APR-2003.
XX
XX 12-AUG-2002; 2002US-00218849.
XX
XX 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX
XX (GETH) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
PI
XX WPI; 2003-644806/61.
DR P-PSDB; ADB83534.
XX
XX New PRO polypeptides and nucleic acids encoding the polypeptides, useful
PT in gene therapy, chromosome identification, tissue typing, or as
PT hybridization probes in chromosome and gene mapping.
XX
XX Claim 2; Fig 43; 315pp; English.
XX
XX The invention describes an isolated PRO (secreted and transmembrane)
CC polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
CC useful for stimulating the proliferation of or gene expression in
CC pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
CC for stimulating the proliferation or differentiation of chondrocyte
CC cells. PRO231, PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide
CC are useful for stimulating the release of tumour necrosis factor (TNF)-
CC alpha from human blood. PRO982, PRO357, PRO725, PRO1306, PRO1419, PRO214,
CC PRO247, PRO337, PRO526, PRO363, PRO531, PRO1083, PRO840, PRO1080,
CC PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
CC PRO1025, PRO1181, PRO1126, PRO1186, PRO1192, PRO1274, PRO1412,
CC PRO1286, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338,
CC PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1567,
CC PRO1887, PRO1928, PRO4344, PRO1801, PRO4333, PRO3543, PRO3444, PRO4322,
CC PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
CC stimulating the proliferation of normal human dermal fibroblasts cells.
CC PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
CC PRO5723, PRO7154, or PRO7425 polypeptide are useful for
CC inhibiting the proliferation of normal human dermal fibroblast cells. PRO
CC polypeptides such as PRO6004, PRO4391, PRO7174, PRO5778, PRO4332, etc.,
CC are useful for detecting the presence of tumour in a mammal which
CC involves comparing the level of expression of the above PRO polypeptides
CC in a test sample of cells taken from the mammal, and a control sample of
CC normal cells of the same cell type, where a higher level of expression of
CC the PRO polypeptides in the test sample as compared to the control sample
CC is indicative of the presence of tumour in the mammal. The tumour is lung

CC tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
CC liver tumour. (I) is useful as molecular weight markers, for tissue
CC typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
CC useful for chromosome and gene mapping or gene therapy. (II) is useful
CC for generating transgenic animals or knock-out animals which are useful
CC screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
CC sport injuries). This sequence encodes a human secreted and transmembrane
CC PRO polypeptide.
XX
SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 9; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAAA 3443
|||||
Db 1347 ATGGCAAA 1391

RESULT 17

ADB80639

ID ADB80639 standard; cDNA; 1434 BP.

XX ADB80639;

DT 04-DEC-2003 (first entry)

XX Novel human secreted and transmembrane protein PRO9905 cDNA.

XX Human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
KW vulnary; antiarthritic; pericyte cell proliferation;
KW pericyte cell differentiation; chondrocyte cell proliferation;
KW chondrocyte cell differentiation; tumour necrosis factor alpha release;
KW (TNF)-alpha release; dermal fibroblast cell proliferation;
KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW liver tumour; tissue typing; chromosome mapping; gene mapping;
KW gene therapy.

XX Homo sapiens.

XX US200308068-A1.

XX 08-MAY-2003.

XX 13-AUG-2002; 2002US-00219481.

XX 01-JUN-2001; 2001WO-US017800.

XX 29-JUN-2001; 2001WO-US021066.

XX 09-APR-2002; 2002US-00119480.

XX (GETH) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;

XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

XX WPI; 2003-657982/62.

XX P-PSDB; ADB80640.

XX One hundred and twenty two nucleic acids encoding PRO polypeptides,
PT useful in gene therapy, chromosome identification, tissue typing, or as
PT hybridization probes in chromosome and gene mapping.

XX Claim 2; Fig 43; 305pp; English.

XX The invention describes an isolated PRO (secreted and transmembrane)
CC polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
CC useful for stimulating the proliferation of or gene expression in
CC pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
CC for stimulating the proliferation or differentiation of chondrocyte
CC cells. PRO231, PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide

CC are useful for stimulating the release of tumour necrosis factor (TNF)-
CC alpha from human blood. PRO982, PRO357, PRO1306, PRO1419, PRO214,
CC PRO247, PRO337, PRO526, PRO363, PRO531, PRO1083, PRO840, PRO1080,
CC PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
CC PRO1025, PRO1181, PRO1126, PRO1186, PRO1192, PRO1244, PRO1274, PRO1412,
CC PRO1286, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338,
CC PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1760, PRO1567,
CC PRO1887, PRO1328, PRO4341, PRO1801, PRO4333, PRO3543, PRO3444, PRO4322,
CC PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
CC stimulating the proliferation of normal human dermal fibroblasts cells.
CC PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
CC PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
CC inhibiting the proliferation of normal human dermal fibroblast cells. PRO
CC polypeptides such as PRO6004, PRO4981, PRO7174, PRO5778, PRO4332, etc.,
CC are useful for detecting the presence of tumour in a mammal which
CC involves comparing the level of expression of the above PRO polypeptides
CC in a test sample of cells taken from the mammal, and a control sample of
CC normal cells of the same cell type, where a higher level of expression of
CC the PRO polypeptides in the test sample as compared to the control sample
CC is indicative of the presence of tumour in the mammal. The tumour is lung
CC tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
CC liver tumour. (I) is useful as molecular weight markers, for tissue
CC typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
CC useful for chromosome and gene mapping or gene therapy. (II) is useful
CC for generating transgenic animals or knock-out animals which are useful
CC screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
CC sport injuries). This sequence encodes a human secreted and transmembrane
CC PRO polypeptide.
XX
SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 9; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAAA 3443
|||||
Db 1347 ATGGCAAA 1391

RESULT 18

ADB73180

ID ADB73180 standard; cDNA; 1434 BP.

XX ADB73180;

XX 04-DEC-2003 (first entry)

XX Novel human secreted and transmembrane protein PRO9905 cDNA.

XX human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
KW vulnary; antiarthritic; pericyte cell proliferation;
KW pericyte cell differentiation; chondrocyte cell proliferation;
KW chondrocyte cell differentiation; tumour necrosis factor alpha release;
KW (TNF)-alpha release; dermal fibroblast cell proliferation;
KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW liver tumour; tissue typing; chromosome mapping; gene mapping;
KW gene therapy.

XX Homo sapiens.

XX US2003096968-A1.

XX 22-MAY-2003.

XX 29-AUG-2002; 2002US-00232223.

XX 01-JUN-2001; 2001WO-US017800.

XX 29-JUN-2001; 2001WO-US021066.

XX 09-APR-2002; 2002US-00119480.

(GETH) GENENTECH INC.

Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PU;
 Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
 WPI; 2003-765525/72.
 P-PSDB; ADB73181.

New isolated PRO polypeptides useful as molecular weight markers in
 protein electrophoresis, useful for tissue typing, and for treating
 arthritis and tumors.

Claim 2; Fig 43; 308pp; English.

The invention describes an isolated PRO (secreted and transmembrane)
 polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
 useful for stimulating the proliferation of or gene expression in
 pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
 for stimulating the proliferation or differentiation of chondrocyte
 cells. PRO231, PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide
 are useful for stimulating the release of tumour necrosis factor (TNF)-
 alpha from human blood. PRO982, PRO357, PRO725, PRO1306, PRO1419, PRO214,
 PRO247, PRO337, PRO526, PRO363, PRO531, PRO1083, PRO840, PRO1080,
 PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
 PRO1025, PRO1181, PRO1126, PRO1186, PRO1192, PRO1274, PRO1412,
 PRO1286, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338,
 PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1760, PRO1567,
 PRO1887, PRO1928, PRO4341, PRO1801, PRO4333, PRO3543, PRO3444, PRO4322,
 PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
 stimulating the proliferation of normal human dermal fibroblasts cells.
 PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
 PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
 inhibiting the proliferation of normal human dermal fibroblast cells. PRO
 polypeptides such as PRO6004, PRO4981, PRO7174, PRO5778, PRO4332, etc.,
 are useful for detecting the presence of tumour in a mammal which
 in a test sample of cells taken from the mammal, and a control sample of
 normal cells of the same cell type, where a higher level of expression of
 the PRO polypeptides in the test sample as compared to the control sample
 is indicative of the presence of tumour in the mammal. The tumour is lung
 tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
 liver tumour. (I) is useful as molecular weight markers, for tissue
 typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
 useful for chromosome and gene mapping or gene therapy. (II) is useful
 for generating transgenic animals or knock-out animals which are useful
 screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
 is useful for treating bone and/or cartilage disorders (e.g., arthritis,
 sport injuries). This sequence encodes a human secreted and transmembrane
 PRO polypeptide.

Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;
 Query Match 1.3%; Score 45; DB 9; Length 1434;
 Best Local Similarity 100.0%; Pred. No. 2.2e-06;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGCAA 3443
 |||
 Db 1347 ATGCAA 1391
 |||

RESULT 19
 ADB78262
 ID ADB78262 standard; cDNA; 1434 BP.

XX
 AC ADB78262;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Novel human secreted and transmembrane protein PRO9905 cDNA.

XX
 KW Human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
 KW vulnary; antiarthritis; pericyte cell proliferation;

KW pericyte cell differentiation; chondrocyte cell proliferation;
 KW chondrocyte cell differentiation; tumour necrosis factor alpha release;
 KW (TNF)-alpha release; dermal fibroblast cell proliferation;
 KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
 KW colon tumour; breast tumour; prostate tumour; rectal tumour;
 KW liver tumour; tissue typing; chromosome mapping; gene mapping;
 KW gene therapy.

XX Homo sapiens.
 OS
 XX
 XX US2003092889-A1.
 PN
 XX
 XX 15-MAY-2003.
 PD
 XX
 XX 13-AUG-2002; 2002US-00219478.
 PF
 XX
 XX 01-JUN-2001; 2001WO-US017800.
 PR
 XX
 XX 29-JUN-2001; 2001WO-US021066.
 PR
 XX
 XX 09-APR-2002; 2002US-00119480.
 PR
 XX
 XX (GETH) GENENTECH INC.
 PA
 XX
 XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PU;
 PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
 PI
 XX
 XX WPI; 2003-765495/72.
 DR
 XX
 XX P-PSDB; ADB78263.
 DR
 XX
 XX New isolated PRO polypeptide useful for tissue typing, gene therapy, as
 PT molecular weight markers in protein electrophoresis, and for treating
 PT arthritis and tumors.

XX
 XX Claim 2; Fig 43; 308pp; English.

XX
 XX The invention describes an isolated PRO (secreted and transmembrane)
 CC polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
 CC useful for stimulating the proliferation of or gene expression in
 CC pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
 CC for stimulating the proliferation or differentiation of chondrocyte
 CC cells. PRO231, PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide
 CC are useful for stimulating the release of tumour necrosis factor (TNF)-
 CC alpha from human blood. PRO982, PRO357, PRO725, PRO1306, PRO1419, PRO214,
 CC PRO247, PRO337, PRO526, PRO363, PRO531, PRO1083, PRO840, PRO1080,
 CC PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
 CC PRO1025, PRO1181, PRO1126, PRO1186, PRO1192, PRO1274, PRO1412,
 CC PRO1286, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338,
 CC PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1760, PRO1567,
 CC PRO1887, PRO1928, PRO4341, PRO1801, PRO4333, PRO3543, PRO3444, PRO4322,
 CC PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
 CC stimulating the proliferation of normal human dermal fibroblasts cells.
 CC PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
 CC PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
 CC inhibiting the proliferation of normal human dermal fibroblast cells. PRO
 CC polypeptides such as PRO6004, PRO4981, PRO7174, PRO5778, PRO4332, etc.,
 CC are useful for detecting the presence of tumour in a mammal which
 CC involves comparing the level of expression of the above PRO polypeptides
 CC in a test sample of cells taken from the mammal, and a control sample of
 CC normal cells of the same cell type, where a higher level of expression of
 CC the PRO polypeptides in the test sample as compared to the control sample
 CC is indicative of the presence of tumour in the mammal. The tumour is lung
 CC tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
 CC liver tumour. (I) is useful as molecular weight markers, for tissue
 CC typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
 CC useful for chromosome and gene mapping or gene therapy. (II) is useful
 CC for generating transgenic animals or knock-out animals which are useful
 CC screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
 CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
 CC sport injuries). This sequence encodes a human secreted and transmembrane
 CC PRO polypeptide.

XX
 XX Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;
 XX
 XX Query Match 1.3%; Score 45; DB 9; Length 1434;
 XX Best Local Similarity 100.0%; Pred. No. 2.2e-06;
 XX Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGCAA 3443
 |||
 Db 1347 ATGCAA 1391
 |||

RESULT 19
 ADB78262
 ID ADB78262 standard; cDNA; 1434 BP.

XX
 AC ADB78262;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Novel human secreted and transmembrane protein PRO9905 cDNA.

XX
 KW Human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
 KW vulnary; antiarthritis; pericyte cell proliferation;

Query Match

1.3%; Score 45; DB 9; Length 1434;

Best Local Similarity 100.0%; Pred. No. 2.2e-06; Mismatches 0; Indels 0; Gaps 0;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAAA 3443
|||
Db 1347 ATGGCAAA 1391

RESULT 20
ADB84910
ID ADB84910 standard; cDNA; 1434 BP.
XX
AC ADB84910;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human PRO polynucleotide #22.
XX
KW Human; PRO; gene; ss; secreted polypeptide; transmembrane polypeptide;
KW tumour; cancer; lung; colon; breast; prostate; rectum; liver;
KW tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell;
KW pericyte cell; dermal fibroblast; bone disorder; cartilage disorder;
KW arthritis; sports injury; cytostatic; antiarthritic.
XX
OS Homo sapiens.
XX
PN US2003073817-A1.
XX
PD 17-APR-2003.
XX
PF 26-AUG-2002; 2002US-00227883.
XX
PR 01-AUG-2000; 2000US-022425P.
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 03-APR-2002; 2002US-00119480.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX
XX WPI: 2003-730024/69.
DR P-PSDB; ADB84911.
XX
PT New PRO polypeptides and nucleic acids encoding the polypeptides, useful
PT e.g. in gene therapy, disease diagnosis, chromosome identification and
PT tissue typing.
XX
PS Claim 2; Fig 43; 314pp; English.
XX
CC The invention relates to human PRO polypeptides (secreted and
CC transmembrane polypeptides) and the PRO polynucleotides encoding them.
CC The PRO polypeptides and polynucleotides are useful as pharmaceuticals,
CC diagnostics, biosensors or bioreactors. They are particularly useful for
CC detecting tumours (e.g. lung tumour, colon tumour, breast tumour,
CC prostate tumour, rectal tumour or liver tumour) in a mammal, for
CC stimulating the release of tumour necrosis factor (TNF)-alpha from human
CC blood, for stimulating the proliferation or differentiation of
CC chondrocyte cells, for stimulating the proliferation of or gene
CC expression in pericyte cells or for stimulating the proliferation of
CC normal human dermal fibroblasts. The PRO nucleic acids are useful as
CC hybridisation probes, in chromosome and gene mapping, in generating
CC antisense RNA and DNA, in preparing PRO polypeptides by recombinant
CC technology, in generating transgenic animals or knock-out animals which
CC may be used in the development and screening of therapeutically useful
CC reagents, in gene therapy, in chromosome identification, as chromosome
CC markers and in generating probes. The PRO polypeptides, or anti-PRO
CC antibodies, are useful for preparing a medicament for treating a
CC condition which is responsive to the PRO polypeptides or anti-PRO
CC antibodies, such as pericyte-associated tumours and bone and/or cartilage
CC disorders (e.g. arthritis, sports injuries), involving inducing the re-
CC differentiation of chondrocytes. The PRO polypeptides are useful as

CC molecular markers for protein electrophoresis, and in tissue typing. This
CC sequence represents a human PRO polynucleotide of the invention.
XX
SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;
Query Match 1.3%; Score 45; DB 10; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3399 ATGGCAAA 3443
|||
Db 1347 ATGGCAAA 1391

RESULT 21
ADB78016
ID ADB78016 standard; cDNA; 1434 BP.
XX
AC ADB78016;
XX
DT 04-DEC-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO9905 cDNA.
XX
KW Human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
KW vulnery; antiarthritic; pericyte cell proliferation;
KW pericyte cell differentiation; chondrocyte cell proliferation;
KW chondrocyte cell differentiation; tumour necrosis factor alpha release;
KW (TNF)-alpha release; dermal fibroblast cell proliferation;
KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW liver tumour; tissue typing; chromosome mapping; gene mapping;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN US2003092886-A1.
XX
PD 15-MAY-2003.
XX
PF 09-AUG-2002; 2002US-00216165.
XX
PR 25-JUL-2000; 2000US-0220607P.
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX
XX (GETH) GENENTECH INC.
XX
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX
XX WPI: 2003-765494/72.
DR P-PSDB; ADB78017.
XX
PT Novel isolated PRO polypeptide useful for tissue typing, gene therapy, as
PT molecular weight markers in protein electrophoresis, for treating
PT arthritis, tumor.
XX
PS Claim 2; Fig 43; 308pp; English.
XX
CC The invention describes an isolated PRO (secreted and transmembrane)
CC polypeptide (1). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
CC useful for stimulating the proliferation of or gene expression in
CC pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
CC for stimulating the proliferation or differentiation of chondrocyte
CC cells. PRO331, PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide
CC are useful for stimulating the release of tumour necrosis factor (TNF)-
CC alpha from human blood. PRO982, PRO357, PRO725, PRO1306, PRO1419, PRO214,
CC PRO247, PRO337, PRO526, PRO363, PRO531, PRO1083, PRO840, PRO1080,
CC PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
CC PRO1025, PRO1181, PRO1126, PRO1186, PRO1192, PRO1244, PRO1412,
CC PRO1286, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338,

CC PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1760, PRO1567,
CC PRO1887, PRO1928, PRO4341, PRO1801, PRO4333, PRO3543, PRO4344, PRO4322,
CC PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
CC stimulating the proliferation of normal human dermal fibroblasts cells.
CC PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
CC PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
CC inhibiting the proliferation of normal human dermal fibroblast cells. PRO
CC polypeptides such as PRO6004, PRO4981, PRO7174, PRO5778, PRO4332, etc.,
CC are useful for detecting the presence of tumour in a mammal which
CC involves comparing the level of expression of the above PRO polypeptides
CC in a test sample of cells taken from the mammal, and a control sample of
CC normal cells of the same cell type, where a higher level of expression of
CC the PRO polypeptides in the test sample as compared to the control sample
CC is indicative of the presence of tumour in the mammal. The tumour is lung
CC tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
CC liver tumour. (I) is useful as molecular weight markers, for tissue
CC typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
CC useful for chromosome and gene mapping or gene therapy. (II) is useful
CC for generating transgenic animals or knock-out animals which are useful
CC screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
CC sport injuries). This sequence encodes a human secreted and transmembrane
CC PRO polypeptide.
XX
SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 10; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGCGCAA 3443
Db 1347 ATGCGCAA 1391

RESULT 22
ADB87082
ID ADB87082 standard; cDNA; 1434 BP.
XX
AC ADB87082;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human PRO polynucleotide #22.
XX
KW Human; PRO; gene; ss; secreted polypeptide; transmembrane polypeptide;
KW tumour; cancer; lung; colon; breast; prostate; rectum; liver;
KW tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell;
KW pericyte cell; dermal fibroblast; bone disorder; cartilage disorder;
KW arthritis; sports injury; cytostatic; antiarthritic.
XX
OS Homo sapiens.
XX
PN US2003088067-A1.
XX
PD 08-MAY-2003.
XX
PF 13-AUG-2002; 2002US-00219479.
XX
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX WPI; 2003-657981/62.
DR P-PSDB; ADB87083.
XX
PT One hundred and twenty two nucleic acids encoding PRO polypeptides, or as
PT useful in gene therapy, chromosome identification, tissue typing, or as

PT hybridization probes in chromosome and gene mapping.
XX
PS Claim 2; Fig 43; 314pp; English.
XX
CC The invention relates to human PRO polypeptides (secreted and
CC transmembrane polypeptides) and the PRO polynucleotides encoding them.
CC The PRO polypeptides and polynucleotides are useful as pharmaceuticals,
CC diagnostics, biosensors or bioreactors. They are particularly useful for
CC detecting tumours (e.g. lung tumour, colon tumour, breast tumour,
CC prostate tumour, rectal tumour or liver tumour) in a mammal, for
CC stimulating the release of tumour necrosis factor (TNF)-alpha from human
CC blood, for stimulating the proliferation or differentiation of
CC chondrocyte cells, for stimulating the proliferation of or gene
CC expression in pericyte cells or for stimulating the proliferation of
CC normal human dermal fibroblasts. The PRO nucleic acids are useful as
CC hybridisation probes, in chromosome and gene mapping, in generating
CC antisense RNA and DNA, in preparing PRO polypeptides by recombinant
CC technology, in generating transgenic animals or knock-out animals which
CC may be used in the development and screening of therapeutically useful
CC reagents, in gene therapy, in chromosome identification, as chromosome
CC markers and in generating probes. The PRO polypeptides, or anti-PRO
CC antibodies, are useful for preparing a medicament for treating a
CC condition which is responsive to the PRO polypeptides or anti-PRO
CC antibodies, such as pericyte-associated tumours and bone and/or cartilage
CC disorders (e.g. arthritis, sports injuries), involving inducing the re-
CC differentiation of chondrocytes. The PRO polypeptides are useful as
CC molecular markers for protein electrophoresis, and in tissue typing. This
CC sequence represents a human PRO polynucleotide of the invention.
XX
SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 10; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGCGCAA 3443
Db 1347 ATGCGCAA 1391

RESULT 23
ADB84664
ID ADB84664 standard; cDNA; 1434 BP.
XX
AC ADB84664;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human PRO polynucleotide #22.
XX
KW Human; PRO; gene; ss; secreted polypeptide; transmembrane polypeptide;
KW tumour; cancer; lung; colon; breast; prostate; rectum; liver;
KW tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell;
KW pericyte cell; dermal fibroblast; bone disorder; cartilage disorder;
KW arthritis; sports injury; cytostatic; antiarthritic.
XX
OS Homo sapiens.
XX
PN US2003092890-A1.
XX
PD 15-MAY-2003.
XX
PF 14-AUG-2002; 2002US-00219536.
XX
PR 28-JUL-1999; 99US-0146222P.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX
PA (GETH) GENENTECH INC.
XX

CC antibodies, are useful for preparing a medicament for treating a
 CC condition which is responsive to the PRO polypeptides or anti-PRO
 CC antibodies, such as pericyte-associated tumours and bone and/or cartilage
 CC disorders (e.g. arthritis, sports injuries), involving inducing the re-
 CC differentiation of chondrocytes. The PRO polypeptides are useful as
 CC molecular markers for protein electrophoresis, and in tissue typing. This
 CC sequence represents a human PRO polynucleotide of the invention.

XX Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 10; Length 1434;
 Best Local Similarity 100.0%; Pred. No. 2.2e-06;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAA 3443
 |||||
 Db 1347 ATGGCAA 1391

RESULT 27

ADCC21762
 ID ADC21762 standard; cDNA; 1434 BP.

XX AC ADC21762;

XX DT 18-DEC-2003 (first entry)

XX DE Human PRO polynucleotide #22.

XX KW Human; PRO; gene; ss; secreted polypeptide; transmembrane polypeptide;
 KW tumour; cancer; lung; colon; breast; prostate; rectum; liver;
 KW tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell;
 KW pericyte cell; dermal fibroblast; bone disorder; cartilage disorder;
 KW arthritis; sports injury; cytostatic; antiarthritic.

XX OS Homo sapiens.

XX PN US2003096969-A1.

XX PD 22-MAY-2003.

XX PF 29-AUG-2002; 2002US-00232225.

XX PR 02-JUN-2000; 2000WO-US015264.
 PR 05-JUN-2000; 2000US-0209832P.
 PR 20-JUN-2000; 2000US-0212901P.
 PR 22-JUN-2000; 2000US-0213807P.
 PR 20-JUL-2000; 2000US-0219556P.
 PR 25-JUL-2000; 2000US-0220585P.
 PR 25-JUL-2000; 2000US-0220605P.
 PR 25-JUL-2000; 2000US-0220607P.
 PR 25-JUL-2000; 2000US-0220624P.
 PR 25-JUL-2000; 2000US-0220638P.
 PR 25-JUL-2000; 2000US-0220664P.
 PR 25-JUL-2000; 2000US-0220666P.
 PR 26-JUL-2000; 2000US-0220893P.
 PR 01-AUG-2000; 2000US-0222425P.
 PR 23-AUG-2000; 2000US-0227133P.
 PR 24-AUG-2000; 2000WO-US023522.
 PR 10-NOV-2000; 2000WO-US023328.
 PR 28-NOV-2000; 2000US-05030873.
 PR 01-DEC-2000; 2000US-0253646P.
 PR 01-DEC-2000; 2000WO-US032678.
 PR 20-DEC-2000; 2000US-00747259.
 PR 28-FEB-2001; 2000WO-US034956.
 PR 25-MAY-2001; 2001WO-US006520.
 PR 01-JUN-2001; 2001WO-US017092.
 PR 29-JUN-2001; 2001WO-US021066.
 PR 09-APR-2002; 2002US-00119480.

XX PA (GETH) GENENTECH INC.

XX

PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
 PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
 XX WPI; 2003-765526/72.
 DR P-PSDB; ADC21763.
 XX

PT Novel isolated PRO polypeptide useful for tissue typing, as molecular
 PT weight markers in protein electrophoresis, for treating arthritis, tumor.

XX Claim 2; Fig 43; 308pp; English.

XX The invention relates to human PRO polypeptides (secreted and
 CC transmembrane polypeptides) and the PRO polynucleotides encoding them.
 CC The PRO polypeptides and polynucleotides are useful as pharmaceuticals,
 CC diagnostics, biosensors or bioreactors. They are particularly useful for
 CC detecting tumours (e.g. lung tumour, colon tumour, breast tumour,
 CC prostate tumour, rectal tumour or liver tumour) in a mammal, for
 CC stimulating the release of tumour necrosis factor (TNF)-alpha from human
 CC blood, for stimulating the proliferation or differentiation of
 CC chondrocyte cells, for stimulating the proliferation of or gene
 CC expression in pericyte cells or for stimulating the proliferation of
 CC normal human dermal fibroblasts. The PRO nucleic acids are useful as
 CC hybridisation probes, in chromosome and gene mapping, in generating
 CC antisense RNA and DNA, in preparing PRO polypeptides by recombinant
 CC technology, in generating transgenic animals or knock-out animals which
 CC may be used in the development and screening of therapeutically useful
 CC reagents, in gene therapy, in chromosome identification, as chromosome
 CC markers and in generating probes. The PRO polypeptides, or anti-PRO
 CC antibodies, are useful for preparing a medicament for treating a
 CC condition which is responsive to the PRO polypeptides or anti-PRO
 CC antibodies, such as pericyte-associated tumours and bone and/or cartilage
 CC disorders (e.g. arthritis, sports injuries), involving inducing the re-
 CC differentiation of chondrocytes. The PRO polypeptides are useful as
 CC molecular markers for protein electrophoresis, and in tissue typing. This
 CC sequence represents a human PRO polynucleotide of the invention.

XX SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 10; Length 1434;
 Best Local Similarity 100.0%; Pred. No. 2.2e-06;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAA 3443
 |||||
 Db 1347 ATGGCAA 1391

RESULT 28

ADC49793

ID ADC49793 standard; cDNA; 1434 BP.

XX AC ADC49793;

XX DT 18-DEC-2003 (first entry)

XX DE Novel human secreted and transmembrane protein PRO9905 cDNA.

XX KW human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
 KW vulnerable; antiarthritic; pericyte cell proliferation;
 KW chondrocyte cell differentiation; chondrocyte cell proliferation;
 KW chondrocyte cell differentiation; tumour necrosis factor alpha release;
 KW (TNF)-alpha release; dermal fibroblast cell proliferation;
 KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
 KW colon tumour; breast tumour; prostate tumour; rectal tumour;
 KW liver tumour; tissue typing; chromosome mapping; gene mapping;
 KW gene therapy.

XX OS Homo sapiens.

XX PN US2003088064-A1.

XX PD 08-MAY-2003.

XX

PF 14-AUG-2002; 2002US-00219075.
XX
PR 25-JUL-2000; 2000US-0220605P.
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX
PA (GETH) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WJ;
XX WPI: 2003-801154/75.
DR P-PSDB; ADC49794.
DR
DR
XX
PT New secreted and transmembrane PRO polypeptide useful for preparing a
PT medicament for treating a condition that is responsive to the PRO
PT polypeptide or anti-PRO antibody, e.g. cancer.
XX
PS Claim 2; SEQ ID NO 43; 314pp; English.
XX
CC The invention describes an isolated PRO (secreted and transmembrane)
CC polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
CC useful for stimulating the proliferation of or gene expression in
CC pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
CC for stimulating the proliferation or differentiation of chondrocyte
CC cells. PRO231, PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide
CC are useful for stimulating the release of tumour necrosis factor (TNF)-
CC alpha from human blood. PRO982, PRO357, PRO725, PRO1306, PRO1419, PRO214,
CC PRO247, PRO337, PRO526, PRO363, PRO831, PRO1083, PRO840, PRO1080,
CC PRO1478, PRO1134, PRO1126, PRO1186, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
CC PRO1025, PRO1181, PRO1286, PRO1347, PRO1305, PRO1273, PRO1274, PRO1412,
CC PRO1286, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338,
CC PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1567,
CC PRO1887, PRO1928, PRO4341, PRO1801, PRO4333, PRO3543, PRO3444, PRO4322,
CC PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
CC stimulating the proliferation of normal human dermal fibroblasts cells.
CC PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4408,
CC PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
CC inhibiting the proliferation of normal human dermal fibroblast cells. PRO
CC polypeptides such as PRO6004, PRO4981, PRO1714, PRO5778, PRO4332, etc.,
CC are useful for detecting the presence of tumour in a mammal which
CC involves comparing the level of expression of the above PRO polypeptides
CC in a test sample of cells taken from the mammal, and a control sample of
CC normal cells of the same cell type, where a higher level of expression of
CC the PRO polypeptides in the test sample as compared to the control sample
CC is indicative of the presence of tumour in the mammal. The tumour is lung
CC tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
CC liver tumour. (I) is useful as molecular weight markers, for tissue
CC typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
CC useful for chromosome and gene mapping or gene therapy. (II) is useful
CC for generating transgenic animals or knock-out animals which are useful
CC screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
CC sport injuries). This sequence encodes a human secreted and transmembrane
CC PRO polypeptide.
XX
SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;
Query Match 1.3%; Score 45; DB 10; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3399 ATGGCAA 3443
|||||
Db 1347 ATGGCAA 1391
|||||
RESULT 29
ADC48992
ID ADC48992 standard; cDNA; 1434 BP.
XX
AC ADC48992;

XX 18-DEC-2003 (first entry)
XX Novel human secreted and transmembrane protein PRO9905 CDNA.
XX
XX human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
XX vulnery; antiarthritic; pericyte cell proliferation;
XX chondrocyte cell differentiation; chondrocyte cell proliferation;
XX (TNF)-alpha release; dermal fibroblast cell proliferation;
XX dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
XX colon tumour; breast tumour; prostate tumour; rectal tumour;
XX liver tumour; tissue typing; chromosome mapping; gene mapping;
XX gene therapy.
XX
XX Homo sapiens.
XX US2003088070-A1.
XX
XX 08-MAY-2003.
XX
XX 28-AUG-2002; 2002US-00230260.
XX
XX 01-JUN-2001; 2001WO-US017800.
XX 29-JUN-2001; 2001WO-US021066.
XX 09-APR-2002; 2002US-00119480.
XX (GETH) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WJ;
XX WPI: 2003-801155/75.
XX P-PSDB; ADC48993.
XX
XX New PRO polypeptides and nucleic acids encoding the polypeptides, useful
XX in gene therapy, chromosome identification, tissue typing, or as
XX hybridization probes in chromosome and gene mapping.
XX
XX Claim 2; SEQ ID NO 43; 315pp; English.
XX
XX The invention describes an isolated PRO (secreted and transmembrane)
XX polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
XX useful for stimulating the proliferation of or gene expression in
XX pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
XX for stimulating the proliferation or differentiation of chondrocyte
XX cells. PRO231, PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide
XX are useful for stimulating the release of tumour necrosis factor (TNF)-
XX alpha from human blood. PRO982, PRO357, PRO725, PRO1306, PRO1419, PRO214,
XX PRO247, PRO337, PRO526, PRO363, PRO831, PRO1083, PRO840, PRO1080,
XX PRO1478, PRO1134, PRO1126, PRO1186, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
XX PRO1025, PRO1181, PRO1286, PRO1347, PRO1305, PRO1273, PRO1274, PRO1412,
XX PRO1286, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338,
XX PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1567,
XX PRO1887, PRO1928, PRO4341, PRO1801, PRO4333, PRO3543, PRO3444, PRO4322,
XX PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
XX stimulating the proliferation of normal human dermal fibroblasts cells.
XX PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4408,
XX PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
XX inhibiting the proliferation of normal human dermal fibroblast cells. PRO
XX polypeptides such as PRO6004, PRO4981, PRO1714, PRO5778, PRO4332, etc.,
XX are useful for detecting the presence of tumour in a mammal which
XX involves comparing the level of expression of the above PRO polypeptides
XX in a test sample of cells taken from the mammal, and a control sample of
XX normal cells of the same cell type, where a higher level of expression of
XX the PRO polypeptides in the test sample as compared to the control sample
XX is indicative of the presence of tumour in the mammal. The tumour is lung
XX tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
XX liver tumour. (I) is useful as molecular weight markers, for tissue
XX typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
XX useful for chromosome and gene mapping or gene therapy. (II) is useful
XX for generating transgenic animals or knock-out animals which are useful
XX screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
XX is useful for treating bone and/or cartilage disorders (e.g., arthritis,
XX sport injuries). This sequence encodes a human secreted and transmembrane
XX PRO polypeptide.

CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
CC sport injuries). This sequence encodes a human secreted and transmembrane
CC PRO polypeptide.

XX
SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 10; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAAA 3443
|||
Db 1347 ATGGCAAA 1391

RESULT 30

ADC49509

ID ADC49509 standard; cDNA; 1434 BP.

AC ADC49509;

XX 18-DEC-2003 (first entry)

XX Novel human secreted and transmembrane protein PRO9905 cDNA.

XX human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
KW vulnery; antiarthritic; pericyte cell proliferation;
KW pericyte cell differentiation; chondrocyte cell proliferation;
KW chondrocyte cell differentiation; tumour necrosis factor alpha release;
KW (TNF)-alpha release; dermal fibroblast cell proliferation;
KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW liver tumour; tissue typing; chromosome mapping; gene mapping;
KW gene therapy.

XX Homo sapiens.

XX US2003088071-A1.

XX 08-MAY-2003.

XX 29-AUG-2002; 2002US-00232231.

XX 01-JUN-2001; 2001WO-US017800.

XX 29-JUN-2001; 2001WO-US021066.

XX 09-APR-2002; 2002US-00119480.

XX (GETH) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;

PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

XX WPI; 2003-801156/75.

XX P-PSDB; ADC49510.

XX New PRO polypeptides and nucleic acids encoding the polypeptides, useful
PT in gene therapy, chromosome identification, tissue typing, or as
PT hybridization probes in chromosome and gene mapping.

XX Claim 2; SEQ ID NO 43; 315pp; English.

XX The invention describes an isolated PRO (secreted and transmembrane)
CC polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
CC useful for stimulating the proliferation of or gene expression in
CC pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
CC for stimulating the proliferation or differentiation of chondrocyte
CC cells. PRO231, PRO357, PRO1275, PRO1155, PRO1306 or PRO1419 polypeptide
CC are useful for stimulating the release of tumour necrosis factor (TNF)-
CC alpha from human blood. PRO982, PRO357, PRO1275, PRO1306, PRO1419, PRO214,
CC PRO247, PRO337, PRO526, PRO363, PRO531, PRO1083, PRO840, PRO1080,
CC PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
CC PRO1025, PRO1181, PRO1126, PRO1186, PRO1192, PRO1274, PRO1412,
CC PRO1286, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338,

CC PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1760, PRO1567,
CC PRO1887, PRO1928, PRO4341, PRO1801, PRO4333, PRO3543, PRO3444, PRO4322,
CC PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
CC stimulating the proliferation of normal human dermal fibroblasts cells.
CC PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
CC PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
CC inhibiting the proliferation of normal human dermal fibroblast cells. PRO
CC polypeptides such as PRO6004, PRO4981, PRO7174, PRO5778, PRO4332, etc.,
CC are useful for detecting the presence of expression of the above PRO polypeptides
CC involves comparing the level of expression of the above PRO polypeptides
CC in a test sample of cells taken from the mammal, and a control sample of
CC normal cells of the same cell type, where a higher level of expression of
CC the PRO polypeptides in the test sample as compared to the control sample
CC is indicative of the presence of tumour in the mammal. The tumour is lung
CC tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
CC liver tumour. (I) is useful as molecular weight markers, for tissue
CC typing or as therapeutic agents. A polynucleotide (II) encoding (I) is
CC useful for chromosome and gene mapping or gene therapy. (II) is useful
CC for generating transgenic animals or knock-out animals which are useful
CC screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
CC sport injuries). This sequence encodes a human secreted and transmembrane
CC PRO polypeptide.

XX Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 10; Length 1434;

Best Local Similarity 100.0%; Pred. No. 2.2e-06;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAAA 3443

|||

Db 1347 ATGGCAAA 1391

Search completed: March 12, 2005, 03:12:39

Job time : 1175 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2005, 02:31:53 ; Search time 388 Seconds
(without alignments)
14519.847 Million cell updates/sec

Title: US-10-627-132-29

Perfect score: 3443

Sequence: 1 ctgcgtgcctgcctgcga.....aaaaaaaaaaaaaaaaaaaa 3443

Scoring table:

OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 25

Total number of hits satisfying chosen parameters: 13194

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Issued Patents NA.*
1: /cgm2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgm2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgm2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgm2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgm2_6/ptodata/1/ina/PTCUS_COMB.seq.*
6: /cgm2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	1.3	1445	3	US-09-697-367-19
2	44	1.3	1445	4	US-09-918-909A-19
3	44	1.3	1530	4	US-09-811-361-19
4	44	1.3	1560	4	US-09-500-495A-5
5	44	1.3	1653	3	US-09-345-469-2
6	44	1.3	1720	3	US-09-227-357-139
7	44	1.3	2311	4	US-09-614-912-91
8	44	1.3	2806	3	US-09-653-839-9
9	44	1.3	2806	4	US-10-202-619-9
10	44	1.3	2852	3	US-09-207-137-2
11	44	1.3	2852	3	US-09-344-441-2
12	44	1.3	3746	4	US-10-160-719A-29
13	44	1.3	3746	4	US-10-160-719A-49
14	44	1.3	3773	4	US-10-160-719A-5
15	44	1.3	3969	4	US-10-160-719A-17
16	44	1.3	3969	4	US-10-160-719A-37
17	44	1.3	5962	6	5386025-5
18	44	1.3	5962	6	5386025-5
19	44	1.3	5975	1	US-08-404-354B-1
20	44	1.3	5975	1	US-08-314-083B-1
21	44	1.3	5975	1	US-08-435-675B-1
22	44	1.3	5975	1	US-08-336-257A-3
23	44	1.3	5975	3	US-08-884-599-1
24	43	1.2	140	4	US-09-621-976-16833
25	43	1.2	141	4	US-09-621-976-16834
26	43	1.2	194	4	US-09-621-976-801
27	43	1.2	286	4	US-09-621-976-17670

28	43	1.2	308	2	US-08-721-488-4	Sequence 4, Appli
29	43	1.2	674	4	US-09-620-405B-465	Sequence 465, App
30	43	1.2	674	4	US-09-433-826B-465	Sequence 465, App
31	43	1.2	674	4	US-09-604-287A-465	Sequence 465, App
32	43	1.2	674	4	US-09-834-759-465	Sequence 465, App
33	43	1.2	674	4	US-09-590-751A-465	Sequence 465, App
34	43	1.2	674	4	US-09-551-621-465	Sequence 465, App
35	43	1.2	1134	3	US-09-248-335-29	Sequence 29, Appl
36	43	1.2	1196	4	US-09-065-040-2	Sequence 2, Appli
37	43	1.2	1525	4	US-09-461-325-110	Sequence 110, App
38	43	1.2	1525	4	US-10-012-542-110	Sequence 110, App
39	43	1.2	1525	4	US-10-115-123-110	Sequence 110, App
40	43	1.2	1559	4	US-09-489-847-42	Sequence 42, Appl
41	43	1.2	1949	4	US-09-461-325-26	Sequence 26, Appl
42	43	1.2	1949	4	US-10-012-542-26	Sequence 26, Appl
43	43	1.2	1949	4	US-10-115-123-26	Sequence 26, Appl
44	43	1.2	2209	1	US-08-514-014-1	Sequence 1, Appli
45	43	1.2	2209	2	US-08-833-823-1	Sequence 1, Appli
46	43	1.2	2280	3	US-08-813-150-1	Sequence 1, Appli
47	43	1.2	2280	4	US-09-546-553-1	Sequence 1, Appli
48	43	1.2	2287	4	US-09-949-016-134	Sequence 134, App
49	43	1.2	2442	4	US-09-575-081B-3	Sequence 3, Appli
50	43	1.2	2950	4	US-09-489-847-92	Sequence 92, Appl
51	43	1.2	2964	4	US-09-578-063-25	Sequence 25, Appl
52	43	1.2	4419	4	US-09-620-312B-187	Sequence 187, App
53	43	1.2	7125	4	US-09-949-016-12359	Sequence 12359, A
54	43	1.2	7156	4	US-09-949-016-16710	Sequence 16710, A
55	42	1.2	133	4	US-09-621-976-8845	Sequence 8845, Ap
56	42	1.2	138	4	US-09-621-976-8112	Sequence 8112, Ap
57	42	1.2	139	4	US-09-621-976-8632	Sequence 8632, Ap
58	42	1.2	150	4	US-09-621-976-8656	Sequence 8656, Ap
59	42	1.2	165	4	US-09-621-976-8127	Sequence 8127, Ap
60	42	1.2	176	4	US-09-621-976-13903	Sequence 13903, A
61	42	1.2	191	4	US-09-621-976-18332	Sequence 18332, A
62	42	1.2	225	4	US-09-621-976-18371	Sequence 18371, A
63	42	1.2	233	4	US-09-621-976-10675	Sequence 10675, A
64	42	1.2	241	4	US-09-621-976-16340	Sequence 16340, A
65	42	1.2	249	4	US-09-621-976-19144	Sequence 19144, A
66	42	1.2	269	4	US-09-621-976-17610	Sequence 17610, A
67	42	1.2	362	4	US-09-621-976-10457	Sequence 10457, A
68	42	1.2	491	4	US-09-311-021-191	Sequence 191, App
69	42	1.2	515	3	US-08-589-028-1	Sequence 1, Appli
70	42	1.2	515	3	US-08-784-582-1	Sequence 1, Appli
71	42	1.2	515	3	US-08-785-271-1	Sequence 1, Appli
72	42	1.2	630	1	US-08-185-414E-1	Sequence 13081, A
73	42	1.2	708	4	US-09-767-13081	Sequence 1878, Ap
74	42	1.2	787	4	US-09-621-976-1878	Sequence 1, Appli
75	42	1.2	806	3	US-08-555-629C-1	Sequence 25, Appl
76	42	1.2	831	4	US-09-904-615-25	Sequence 1, Appli
77	42	1.2	857	1	US-08-308-883-1	Sequence 1, Appli
78	42	1.2	857	1	US-08-730-163-1	Sequence 1, Appli
79	42	1.2	857	3	US-08-256-799-1	Sequence 1, Appli
80	42	1.2	857	3	US-08-462-437-1	Sequence 1, Appli
81	42	1.2	941	4	US-09-205-258-186	Sequence 186, App
82	42	1.2	971	4	US-09-800-729-49	Sequence 49, Appl
83	42	1.2	985	4	US-09-322-409-25	Sequence 25, Appl
84	42	1.2	985	4	US-09-322-409-27	Sequence 27, Appl
85	42	1.2	985	4	US-09-451-527-25	Sequence 25, Appl
86	42	1.2	985	4	US-09-451-527-27	Sequence 27, Appl
87	42	1.2	1023	1	US-08-252-266B-16	Sequence 16, Appl
88	42	1.2	1036	4	US-09-205-258-86	Sequence 86, Appl
89	42	1.2	1050	4	US-09-482-273-58	Sequence 58, Appl
90	42	1.2	1066	1	US-08-157-101A-4	Sequence 4, Appli
91	42	1.2	1181	3	US-09-149-476-310	Sequence 310, App
92	42	1.2	1212	3	US-09-149-476-186	Sequence 186, App
93	42	1.2	1307	2	US-08-960-022-17	Sequence 17, Appl
94	42	1.2	1358	4	US-09-949-016-463	Sequence 463, App
95	42	1.2	1359	3	US-09-387-574-11	Sequence 11, Appl
96	42	1.2	1359	3	US-09-668-096-11	Sequence 11, Appl
97	42	1.2	1405	4	US-09-668-097A-15	Sequence 15, Appl
98	42	1.2	1459	4	US-09-537-654-3	Sequence 3, Appli
99	42	1.2	1503	3	US-08-923-454A-5	Sequence 5, Appli
100	42	1.2	1534	1	US-08-300-903A-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1

US-09-697-367-19
; Sequence 19, Application US/09697367
; Patent No. 6323015
; GENERAL INFORMATION:
; APPLICANT: Orozco Jr., Emil M.
; APPLICANT: Caiami, Perry G.
; APPLICANT: Weng, Zude
; APPLICANT: Tarczynski, Mitchell
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE
; FILE REFERENCE: BB1166 US NA
; CURRENT APPLICATION NUMBER: US/09/697,367
; CURRENT FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/084,529
; PRIOR FILING DATE: 1998-MAY-07
; PRIOR APPLICATION NUMBER: PCT/US99/09865
; PRIOR FILING DATE: 1999-MAY-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 1445
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-697-367-19

Query Match 1.3%; Score 44; DB 3; Length 1445;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3400 TGGCAA 3443
|||
DB 1392 TGGCAA 1435

RESULT 2

US-09-918-909A-19
; Sequence 19, Application US/09918909A
; Patent No. 6756218
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Caiami, Perry G.
; APPLICANT: Orozco Jr., Emil M.
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE
; FILE REFERENCE: BB1166 US CIP
; CURRENT APPLICATION NUMBER: US/09/918,909A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/084,529
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: PCT/US99/09865
; PRIOR FILING DATE: 1999-05-06
; PRIOR APPLICATION NUMBER: 09/697,367
; PRIOR FILING DATE: 2000-10-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 1445
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-918-909A-19

Query Match 1.3%; Score 44; DB 4; Length 1445;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3400 TGGCAA 3443
|||
DB 1392 TGGCAA 1435

RESULT 3

US-09-811-361-19
; Sequence 19, Application US/09811361
; Patent No. 6730821
; GENERAL INFORMATION:
; APPLICANT: Guenther, Catherine
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING
; TITLE OF INVENTION: RETINA-SPECIFIC NUCLEAR RECEPTOR GENE DISRUPTIONS
; FILE REFERENCE: R-125
; CURRENT APPLICATION NUMBER: US/09/811,361
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: US 60/190,348
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 1530
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-811-361-19

Query Match 1.3%; Score 44; DB 4; Length 1530;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3400 TGGCAA 3443
|||
DB 1478 TGGCAA 1521

RESULT 4

US-09-500-495A-5
; Sequence 5, Application US/09500495A
; Patent No. 6812382
; GENERAL INFORMATION:
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Hitz, William D.
; APPLICANT: Lohman, Karin N.
; APPLICANT: Miao, Gou-Hau
; APPLICANT: Orozco, Emil M.
; TITLE OF INVENTION: Plant Nucleic Acids Encoding Chloroplast Division Proteins
; FILE REFERENCE: BB1338 US NA
; CURRENT APPLICATION NUMBER: US/09/500,495A
; CURRENT FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: 60/119,419
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-500-495A-5

Query Match 1.3%; Score 44; DB 4; Length 1560;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3400 TGGCAA 3443
|||
DB 1491 TGGCAA 1534

RESULT 5

US-09-345-469-2
; Sequence 2, Application US/09345469
; Patent No. 6369210
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: 22012, A No. 6369210e1 Human Carboxypeptidase

FILE REFERENCE: 5800-38
CURRENT APPLICATION NUMBER: US/09/345,469
CURRENT FILING DATE: 1993-06-30
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1653
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (19)...(1494)
US-09-345-469-2

Query Match 1.3%; Score 44; DB 3; Length 1653;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3400 TGGCAA 3443
Db 1596 TGGCAA 1639

RESULT 6

US-09-227-357-139
Sequence 139, Application US/09227357
Patent No. 6342581
GENERAL INFORMATION:
APPLICANT: Fischer et al.
TITLE OF INVENTION: 123 Human Secreted Proteins
FILE REFERENCE: P2010F1
CURRENT APPLICATION NUMBER: US/09/227,357
CURRENT FILING DATE: 1999-01-08
EARLIER APPLICATION NUMBER: PCT/US98/13684
EARLIER FILING DATE: 1998-07-07
EARLIER APPLICATION NUMBER: 60/051,926
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,793
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,925
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,929
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,932
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,916
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,930
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18

EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 139
LENGTH: 1720
TYPE: DNA
ORGANISM: Homo sapiens
US-09-227-357-139

Query Match 1.3%; Score 44; DB 3; Length 1720;

Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3400 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 1670 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713

RESULT 7

US-09-614-912-91
Sequence 91, Application US/09614912
Patent No. 6677502
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Rafalski, Antoni
APPLICANT: Orozco, Buddy
APPLICANT: Miao, Gou-Hau
APPLICANT: Famodu, Omolayo O.
APPLICANT: Lee, Jian Ming
APPLICANT: Sakai, Hajime
APPLICANT: Weng, Zude
APPLICANT: Caimi, Perry G
APPLICANT: Anderson, Shawn
TITLE OF INVENTION: Plant Metabolism Genes
FILE REFERENCE: BB1378 US NA
CURRENT APPLICATION NUMBER: US/09/614,912
CURRENT FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: 60/143,401
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/143,412
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/146,650
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 60/170,906
PRIOR FILING DATE: 1999-12-15
PRIOR APPLICATION NUMBER: 60/172,959
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/172,946

```

; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 91
; LENGTH: 2311
; TYPE: DNA
; ORGANISM: Glycine max
US-09-614-912-91

```

```
Query Match      1.3%; Score 44; DB 4; Length 2311;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	3400	TGGCAAAAAAAAAAAAAAAAAAAAAAAAAA	3443
Db	2255	TGGCAAAAAAAAAAAAAAAAAAAAAAAAAA	2298

```

RESULT 8
US-09-653-839-9
; Sequence 9, Application US/09653839
; Patent No. 6433153
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6433153el Human Calcium Dependent Proteases
; TITLE OF INVENTION: and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0038-USA
; CURRENT APPLICATION NUMBER: US/09/653,839
; CURRENT FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 60/152,057
; PRIOR FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2806
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-653-839-9

```

```

Query Match      1.3%; Score 44; DB 3; Length 2806;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3400 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 2709 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2752

```

```

RESULT 9
US-10-202-619-9
; Sequence 9, Application US/10202619
; Patent No. 6716614
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6716614e1 Human Calcium Dependent Proteases,
; TITLE OF INVENTION: Polynucleotides Encoding the Same, and Uses Thereof
; FILE REFERENCE: LEX-0391-USA
; CURRENT APPLICATION NUMBER: US/10/202,619
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/152,057
; PRIOR FILING DATE: 1999-09-02
; PRIOR APPLICATION NUMBER: US 09/653,839
; PRIOR FILING DATE: 2000-09-01

```

```

; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2806
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-202-619-9

```

Query Match	1.3%	Score 44;	DB 4;	Length 2806;
Best Local Similarity	100.0%;	Pred. No.	1.3e-08;	
Matches 44;	Conservative	0;	Mismatches 0;	Indels
Qy	3400	TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAAA	3443
Dy	2709	TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAAA	2752

RESULT 10
 US-09-027-137-2
 Sequence 2, Application US/09027137
 Patent No. 6013450
 GENERAL INFORMATION:
 APPLICANT: Hillman, Jennifer L.
 APPLICANT: Corley, Neil C.
 APPLICANT: Yue, Henry
 TITLE OF INVENTION: CAF1-RELATED PROTEIN
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Dr.
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/027,137
 FILING DATE: Filed Herewith
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0476 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-855-0555
 TELEFAX: 650-845-4166
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2852 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: PROSNOT16
 CLONE: 2229466
 US-09-027-137-2

Query Match	1.3%;	Score 44;	DB 3;	Length 2852;
Best Local Similarity	100.0%;	Pred. No. 1.2e-08;		
Matches 44;	Conservative 0;	Mismatches 0;	Indels	
Qy	3400	TGGCAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	3443	
Db	2497	TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2540	

RESULT 11
US-09-344-441-2
; Sequence 2, Application US/09344441
; Patent No. 6376651
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; Corley, Neil C.
; Yue, Henry
; TITLE OF INVENTION: CAP1-RELATED PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/344,441
; FILING DATE: 20-Feb-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/027,137
; FILING DATE: 1998-02-20
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0476 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2852 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT16
; CLONE: 2229466
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-344-441-2

Query Match 1.3%; Score 44; DB 3; Length 2852;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3400 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 2497 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2540

RESULT 12
US-10-160-719A-29
; Sequence 29, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17

; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 3746
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (321)...(3551)
US-10-160-719A-29
Query Match 1.3%; Score 44; DB 4; Length 3746;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3038 CTCTTCGGCAAGCTCTTCTTCGGCTTCTGGGTATCGTCCACCT 3081
Db 3342 CTCTTCGGCAAGCTCTTCTTCGGCTTCTGGGTATCGTCCACCT 3385

RESULT 13
US-10-160-719A-49
; Sequence 49, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 3746
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (321)...(3449)
US-10-160-719A-49

Query Match 1.3%; Score 44; DB 4; Length 3746;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3038 CTCTTCGGCAAGCTCTTCTTCGGCTTCTGGGTATCGTCCACCT 3081
Db 3342 CTCTTCGGCAAGCTCTTCTTCGGCTTCTGGGTATCGTCCACCT 3385

RESULT 14
US-10-160-719A-5
; Sequence 5, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 1998-08-17

```

; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3773
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (338)...(3568)
US-10-160-719A-5

Query Match      1.3%; Score 44; DB 4; Length 3773;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3038 CTCTTCGGCAAGCTCTTCTTCGCGCTTCTGGGTCTCATCGTCCACCT 3081
      |||||||
Db 3359 CTCTTCGGCAAGCTCTTCTTCGCGCTTCTGGGTCTCATCGTCCACCT 3402
      |||||||

RESULT 15
US-10-160-719A-17
; Sequence 17, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwardal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3969
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (144)...(3401)
US-10-160-719A-17

Query Match      1.3%; Score 44; DB 4; Length 3969;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 299 TACGAGCGCCGGGAGGGCAGCAGAACTGCCCCCAGTGCAAGAC 342
      |||||||
Db 357 TACGAGCGCCGGGAGGGCAGCAGAACTGCCCCCAGTGCAAGAC 400
      |||||||

RESULT 16
US-10-160-719A-37
; Sequence 37, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwardal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses

```

```

1  APPLICANT: Harpold, Michael Miller
2  APPLICANT: Schwartz, Arnold
3  APPLICANT: Brenner, Robert
4  TITLE OF INVENTION: CALCIUM CHANNEL COMPOSITIONS AND METHODS
5  NUMBER OF SEQUENCES: 3
6  CORRESPONDENCE ADDRESS:
7  ADDRESSEE: Brown, Martin, Haller & McClain
8  STREET: 1660 Union Street
9  CITY: San Diego
10 STATE: CA
11 COUNTRY: USA
12 ZIP: 92101-2926
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: Diskette
15 COMPUTER: IBM Compatible
16 OPERATING SYSTEM: DOS
17 SOFTWARE: FastSEQ Version 1.5
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER: US/08/314,083B
20 FILING DATE: 28-SEPT-1994
21 CLASSIFICATION: 435
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER: US 07/914,231
24 FILING DATE: 13-JUL-1992
25 PRIOR APPLICATION DATA:
26 APPLICATION NUMBER: US 07/603,751
27 FILING DATE: 08-NOV-1990
28 ATTORNEY/AGENT INFORMATION:
29 NAME: Seidman, Stephanie L
30 REGISTRATION NUMBER: 33,779
31 REFERENCE/DOCKET NUMBER: 6362-53191
32 TELECOMMUNICATION INFORMATION:
33 TELEPHONE: 619-238-0999
34 TELEFAX: 619-238-0062
35 TELEX:
36 INFORMATION FOR SEQ ID NO: 1:
37 SEQUENCE CHARACTERISTICS:
38 LENGTH: 5975 base pairs
39 TYPE: nucleic acid
40 STRANDEDNESS: single
41 TOPOLOGY: linear
42 MOLECULE TYPE: cdna
43 FEATURE:
44 NAME/KEY: Coding Sequence
45 LOCATION: 79...5700
46 OTHER INFORMATION:
47 US-08-314-083B-1
48
49 Query Match 1.3%; Score 44; DB 1; Length 5975;
50 Best Local Similarity 100.0%; Pred.No. 1.2e-08;
51 Matches 44; Conservative 0; Mismatches 0; Indels 0
52
53 QY 3400 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
54 |||
55 Db 5928 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 5971
56 |||
57
58 RESULT 21
59 US-08-435-675B-1
60 Sequence 1, Application US/08435675B
61 Patent No. 5710250
62 GENERAL INFORMATION:
63 APPLICANT: Ellis, Steven Bradley
64 APPLICANT: Williams, Mark E.
65 APPLICANT: Harpold, Michael Miller
66 APPLICANT: Schwartz, Arnold
67 APPLICANT: Brenner, Robert
68 TITLE OF INVENTION: CALCIUM CHANNEL COMPOSITIONS AND METHODS
69 NUMBER OF SEQUENCES: 6
70 CORRESPONDENCE ADDRESS:
71 ADDRESSEE: Brown, Martin, Haller & McClain
72 STREET: 1660 Union Street
73 CITY: San Diego

```

STATE: CA
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/435,675B
FILING DATE: 05-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/314,083
FILING DATE: 28-SEP-1994
APPLICATION NUMBER: US 07/914,231
FILING DATE: 13-JUL-1992
APPLICATION NUMBER: US 07/603,751
FILING DATE: 08-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-53193
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-238-0999
TELEFAX: 619-238-0062
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5975 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 79...5700
OTHER INFORMATION:
US-08-435-675B-1

Query Match 1.3%; Score 44; DB 1; Length 5975;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3400 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|||||
DB 5928 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 5971

RESULT 22
US-08-336-257A-3
Sequence 3, Application US/08336257A
Patent No. 5726035
GENERAL INFORMATION:
APPLICANT: Jay, Scott D
APPLICANT: Ellis, Steven B.
APPLICANT: Harpold, Michael M.
APPLICANT: Campbell, Kevin P.
TITLE OF INVENTION: CALCIUM CHANNEL COMPOSITIONS AND METHODS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/336,257A
FILING DATE: 07-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 54898
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 238-0999
TELEFAX: (619) 238-0062
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 5975 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: Genomic DNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 79...5700
OTHER INFORMATION: \product= "Alpha-1 subunit of animal calcium
channel"
US-08-336-257A-3

Query Match 1.3%; Score 44; DB 1; Length 5975;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3400 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|||||
DB 5928 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 5971

RESULT 23
US-08-884-599-1
Sequence 1, Application US/08884599
Patent No. 6013474
GENERAL INFORMATION:
APPLICANT: Ellis, Steven Bradley
APPLICANT: Williams, Mark E.
APPLICANT: Harpold, Michael Miller
APPLICANT: Schwartz, Arnold
APPLICANT: Brenner, Robert
TITLE OF INVENTION: CALCIUM CHANNEL COMPOSITIONS AND METHODS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/884,599
FILING DATE: 27-JUNE-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/314,083
FILING DATE: 28-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/914,231
FILING DATE: 13-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 08-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-53191B

TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-238-0999
TELEFAX: 619-238-0062
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5975 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 79...5700
OTHER INFORMATION:
US-08-884-599-1

Query Match 1.3%; Score 44; DB 3; Length 5975;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3400 TGGCAA 3443
Db 5928 TGGCAA 5971

RESULT 24

US-09-621-976-16833
Sequence 16833, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 16833
LENGTH: 140
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-16833

Query Match 1.2%; Score 43; DB 4; Length 140;
Best Local Similarity 100.0%; Pred. No. 4.1e-08;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3401 GGCAA 3443
Db 98 GGCAA 140

RESULT 25

US-09-621-976-16834
Sequence 16834, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 16834
LENGTH: 141
TYPE: DNA
ORGANISM: Homo sapiens

US-09-621-976-16834

Query Match 1.2%; Score 43; DB 4; Length 141;
Best Local Similarity 100.0%; Pred. No. 4.1e-08;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3401 GGCAA 3443
Db 98 GGCAA 140

RESULT 26

US-09-621-976-801
Sequence 801, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 801
LENGTH: 194
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 20..193
US-09-621-976-801

Query Match 1.2%; Score 43; DB 4; Length 194;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3401 GGCAA 3443
Db 98 GGCAA 140

RESULT 27

US-09-621-976-17670
Sequence 17670, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 17670
LENGTH: 286
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-17670

Query Match 1.2%; Score 43; DB 4; Length 286;
Best Local Similarity 100.0%; Pred. No. 3.9e-08;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3401 GGCAA 3443
Db 240 GGCAA 282

RESULT 28

```
US-08-721-488-4
; Sequence 4, Application US/08721488
; Patent No. 5965388
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: LaVallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; APPLICANT: Bowman, Michael
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/721,488
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 308 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-721-488-4
```

```
Query Match 1.2%; Score 43; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 3.9e-08;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3401 GGCACAAAAA 3443
|||
Db 266 GGCACAAAAA 308
```

```
RESULT 29
US-09-620-405B-465
; Sequence 465, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 465
; LENGTH: 674
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-620-405B-465
```

```
Query Match 1.2%; Score 43; DB 4; Length 674;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3401 GGCACAAAAA 3443
|||
Db 547 GGCACAAAAA 589
```

RESULT 30

```
US-09-433-826B-465
; Sequence 465, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C4
; CURRENT APPLICATION NUMBER: US/09/433,826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 465
; LENGTH: 674
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-433-826B-465
```

```
Query Match 1.2%; Score 43; DB 4; Length 674;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3401 GGCACAAAAA 3443
|||
Db 547 GGCACAAAAA 589
```

```
Search completed: March 12, 2005, 07:57:19
Job time : 392 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2005, 05:53:54 ; Search time 1222 Seconds
(without alignments)
16760.046 Million cell updates/sec

Title: US-10-627-132-29

Perfect score: 3443

Sequence: 1 cfcgctgcctcctcgcga.....aaaaaaaaaaaaaaaa 3443

Scoring table: OLIGO NUC
Gapop 60-0, Gapext 60.0

Searched: 5537552 seqs, 2974263231 residues

Word size : 25

Total number of hits satisfying chosen parameters: 131458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3443	100.0	3443	17	US-10-627-132-29
2	3028	87.9	3028	16	US-10-209-059-29
3	2583	75.0	2687	17	US-10-425-114-833
4	173	5.0	2830	18	US-10-425-115-13789
5	173	5.0	3589	17	US-10-425-114-24930
6	124	3.6	335	18	US-10-425-115-179459
7	116	3.4	999	18	US-10-767-701-10062
8	110	3.2	633	18	US-10-425-115-10141
9	83	2.4	3448	18	US-10-437-963-15093
10	59	1.7	549	18	US-10-767-701-29317
11	58	1.7	488	18	US-10-425-115-13788
					Sequence 29, Appl
					Sequence 29, Appl
					Sequence 833, Appl
					Sequence 139789, A
					Sequence 24930, A
					Sequence 179459, A
					Sequence 10062, A
					Sequence 10141, A
					Sequence 15093, A
					Sequence 29317, A
					Sequence 139788, A

85 45 1.3 1434 15 US-10-219-467-43 Sequence 43, Appl
86 45 1.3 1434 15 US-10-219-469-43 Sequence 43, Appl
87 45 1.3 1434 15 US-10-219-471-43 Sequence 43, Appl
88 45 1.3 1434 15 US-10-219-473-43 Sequence 43, Appl
89 45 1.3 1434 15 US-10-219-476-43 Sequence 43, Appl
90 45 1.3 1434 15 US-10-219-482-43 Sequence 43, Appl
91 45 1.3 1434 15 US-10-227-874-43 Sequence 43, Appl
92 45 1.3 1434 15 US-10-227-876-43 Sequence 43, Appl
93 45 1.3 1434 15 US-10-227-878-43 Sequence 43, Appl
94 45 1.3 1434 15 US-10-229-974-43 Sequence 43, Appl
95 45 1.3 1434 15 US-10-230-024-43 Sequence 43, Appl
96 45 1.3 1434 15 US-10-230-113-43 Sequence 43, Appl
97 45 1.3 1434 15 US-10-230-183-43 Sequence 43, Appl
98 45 1.3 1434 15 US-10-230-234-43 Sequence 43, Appl
99 45 1.3 1434 15 US-10-230-306-43 Sequence 43, Appl
100 45 1.3 1434 15 US-10-230-426-43 Sequence 43, Appl

ALIGNMENTS

RESULT 1
US-10-627-132-29
; Sequence 29, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 3443
; TYPE: DNA
; ORGANISM: Zea mays
US-10-627-132-29

Query Match 100.0%; Score 3443; DB 17; Length 3443;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCGTGGCCCTCGCNAATCGGAATCTGTGAGACACCTGAGGGTGGAGGCCGAG 60
Db 1 CTGCGTGGCCCTCGCNAATCGGAATCTGTGAGACACCTGAGGGTGGAGGCCGAG 60
Qy 61 AGCTAGCTAGCACGCGGCTCCGCGCGAGATGGAGGCCAGCGCGGCTGTGSCCG 120
Db 61 AGCTAGCTAGCACGCGGCTCCGCGCGAGATGGAGGCCAGCGCGGCTGTGSCCG 120
Qy 121 CTCGCAACACCGGAAACGAGTGTGTGATTCGGGGCCACGAGGACCCCAAGCCGTGCG 180
Db 121 CTCGCAACACCGGAAACGAGTGTGTGATTCGGGGCCACGAGGACCCCAAGCCGTGCG 180
Qy 181 GCGCTGAGCGGCGAGGTGTGAGATATCGCGGCGACGAGTGGCTTACCGTGAACGG 240
Db 181 GCGCTGAGCGGCGAGGTGTGAGATATCGCGGCGACGAGTGGCTTACCGTGAACGG 240
Qy 241 CGACCTCTTCGTCGCTCGCAACAGTGGGCTTCCCGGTGTGCGGCGCTGTACAGTA 300
Db 241 CGACCTCTTCGTCGCTCGCAACAGTGGGCTTCCCGGTGTGCGGCGCTGTACAGTA 300

301 CGAGCGCCGGGAGGGGACGCGAAGCTGCCCCCAAGTGCAGAGCGCGCTACAAGCGCCTCAA 360
301 CGAGCGCCGGGAGGGGACGCGAAGCTGCCCCCAAGTGCAGAGCGCGCTACAAGCGCCTCAA 360
361 GGGGAGCCCGAGGGTTCGCGGGAGCGATGACGAGGAGGACATCGAGCTCGAGCACCA 420
361 GGGGAGCCCGAGGGTTCGCGGGAGCGATGACGAGGAGGACATCGAGCTCGAGCACCA 420
421 GTTCAACATCTGACGACGAGAATCAGCAGAGGAGCTGGAGGGCAACATCAGACAAAGCCCA 480
421 GTTCAACATCTGACGACGAGAATCAGCAGAGGAGCTGGAGGGCAACATCAGACAAAGCCCA 480
481 GATCACCGAGGCGATGCTGCAACGAGGATGAGCTACGAGGAGGGCCCCGACGACGCGCA 540
481 GATCACCGAGGCGATGCTGCAACGAGGATGAGCTACGAGGAGGGCCCCGACGACGCGCA 540
541 CGGCAACACACCCCGCAGATCCCGCCCATCATACCCGGCTCCCGTCCGTCGCGGTGAG 600
541 CGGCAACACACCCCGCAGATCCCGCCCATCATACCCGGCTCCCGTCCGTCGCGGTGAG 600
601 CGGTGAGTTTCCGATTACCAACGGGTATGCGCACCGGAGGTCTCGTCTTCCCTGCACAA 660
601 CGGTGAGTTTCCGATTACCAACGGGTATGCGCACCGGAGGTCTCGTCTTCCCTGCACAA 660
661 GCGCATCCATCCGTACCCCTGTCTGAGCCAGGGAGTGCNAAGTGGGACGAGAAGAAAGA 720
661 GCGCATCCATCCGTACCCCTGTCTGAGCCAGGGAGTGCNAAGTGGGACGAGAAGAAAGA 720
721 AGTGAGCTGGAGGAGAGATGGAACGCTGGCACTGAAACGAGGCGAGGCGAGCCGCT 840
721 AGTGAGCTGGAGGAGAGATGGAACGCTGGCACTGAAACGAGGCGAGGCGAGCCGCT 840
841 GTGAGGAAAGGTGTGATCGCTCGAGCAAGGTGAACCCGTACCGGATGGTGAATCGTGGT 900
841 GTGAGGAAAGGTGTGATCGCTCGAGCAAGGTGAACCCGTACCGGATGGTGAATCGTGGT 900
901 GGGTCTCGTTGTGCTCGCCTTCTTCCGCGTACCGTATCTCTGCAACCCCGTCCCGACGC 960
901 GGGTCTCGTTGTGCTCGCCTTCTTCCGCGTACCGTATCTCTGCAACCCCGTCCCGACGC 960
961 CATCGGGCTGTGGCTCGTCTCCATCATCTGCGAGATCTGGTTCGCCATCTCTCTGGATCCT 1020
961 CATCGGGCTGTGGCTCGTCTCCATCATCTGCGAGATCTGGTTCGCCATCTCTCTGGATCCT 1020
1021 CGACAGTTTCCCAAGTGGTTCGCCATCGACCGGAGAGCTACCTCGACCGGCTCTCCCT 1080
1021 CGACAGTTTCCCAAGTGGTTCGCCATCGACCGGAGAGCTACCTCGACCGGCTCTCCCT 1080
1081 CAGGTACGAGAGGAGGAGCGCTCGTCTGCTGCGGGTGGACCTGTTGTTGAGCAC 1140
1081 CAGGTACGAGAGGAGGAGCGCTCGTCTGCTGCGGGTGGACCTGTTGTTGAGCAC 1140
1141 GGTGACACCGCTCAAGGAGCCCGCTGGTGAACCGCAACACCGTGTCTCTCATCTCTCGC 1200
1141 GGTGACACCGCTCAAGGAGCCCGCTGGTGAACCGCAACACCGTGTCTCTCATCTCTCGC 1200
1201 CGTAGACTACCCCGTGGACAAAGTCTCTGCTACGTTCCGACGAGCGGCGTCTGATGCT 1260
1201 CGTAGACTACCCCGTGGACAAAGTCTCTGCTACGTTCCGACGAGCGGCGTCTGATGCT 1260
1261 GAGCTTCGAGTTCGAGTTCGAGAGCGGCGAGTTCGCGCGCAAGTGGGTGCCCTTCTGCAA 1320
1261 GAGCTTCGAGTTCGAGTTCGAGAGCGGCGAGTTCGCGCGCAAGTGGGTGCCCTTCTGCAA 1320
1321 GAAGTTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGT 1380
1321 GAAGTTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGT 1380
1381 CAAAGCAAGGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGT 1440

[illegible]

Db	2461	CCGCTGGCGCTGGGTCCTGCAGATCTTTCTTTCAGCCGGCACAGGCCCTCCTGTCTACCG	2522
QY	2521	CTACAAGAACGGCAACTCTAACTGGTGAGCGCTTCGCCTACATCAAACACCACCATCTA	2580
Db	2521	CTACAAGAACGGCAACTCTAACTGGTGAGCGCTTCGCCTACATCAAACACCACCATCTA	2580
QY	2581	CCCCTTCACTCGCTCCCGCTGTCTCGCTACTCGAACCCTCCCGCGCTTCGCTCCTCAC	2640
Db	2581	CCCCTTCACTCGCTCCCGCTGTCTCGCTACTCGAACCCTCCCGCGCTTCGCTCCTCAC	2640
QY	2641	CGGCAAGTTTCATCATGCGCGTGAAGCAAGTTTCGCGACGCTCTTCTTCATCGCCCTCTT	2700
Db	2641	CGGCAAGTTTCATCATGCGCGTGAAGCAAGTTTCGCGACGCTCTTCTTCATCGCCCTCTT	2700
QY	2701	CATGTCATCTTTCGCAACGGGCATCTCTGGAGATGCGGTGGAGCGGGGTGAGATCGAGGA	2760
Db	2701	CATGTCATCTTTCGCAACGGGCATCTCTGGAGATGCGGTGGAGCGGGGTGAGATCGAGGA	2760
QY	2761	GTTGTTGAGGAACGAGCAGATTCTGGGTTCATCGCGCGGTGTTCGCGCATCTCTTCGCCGT	2820
Db	2761	GTTGTTGAGGAACGAGCAGATTCTGGGTTCATCGCGCGGTGTTCGCGCATCTCTTCGCCGT	2820
QY	2821	CGTGCAGGGCTGCTCAAAGTCTCTCGCGGGATCGACCACTTCACCGTCACTCTCCAA	2880
Db	2821	CGTGCAGGGCTGCTCAAAGTCTCTCGCGGGATCGACCACTTCACCGTCACTCTCCAA	2880
QY	2881	GGCCACCGGCGAGGAGCAACAGTTTCGCCGAGCTCTACGCCCTTCAAGTGGACCAACGCT	2940
Db	2881	GGCCACCGGCGAGGAGCAACAGTTTCGCCGAGCTCTACGCCCTTCAAGTGGACCAACGCT	2940
QY	2941	CCTCATCCCCGCCACACACGCTGCTCATATTAAGTTCATCGCGGTGTTGCGCGGATCTC	3000
Db	2941	CCTCATCCCCGCCACACACGCTGCTCATATTAAGTTCATCGCGGTGTTGCGCGGATCTC	3000
QY	3001	CGACGCCATCAAACAAGGGTACCAGTCTCTGGGGCCCCCTCTTCGGCAAGCTCTTCTTCGC	3060
Db	3001	CGACGCCATCAAACAAGGGTACCAGTCTCTGGGGCCCCCTCTTCGGCAAGCTCTTCTTCGC	3060
QY	3061	CTTCTGGGTTCATCGTCCACCTTACCCGTTCTCCTCAAGGGGCTCATGGGCGCCAGAACAG	3120
Db	3061	CTTCTGGGTTCATCGTCCACCTTACCCGTTCTCCTCAAGGGGCTCATGGGCGCCAGAACAG	3120
QY	3121	GACGCCACCGTTGTTGTTCATCTGGTCCATTCTGCTGGGCTCCATCTTCTCCCTGCTCTG	3180
Db	3121	GACGCCACCGTTGTTGTTCATCTGGTCCATTCTGCTGGGCTCCATCTTCTCCCTGCTCTG	3180
QY	3181	GGTCAGGATCGACCCCTTTTCATCTGTACAGGACCAAGGCGCCGGAGCTCAGGCGAGTGGCAT	3240
Db	3181	GGTCAGGATCGACCCCTTTTCATCTGTACAGGACCAAGGCGCCGGAGCTCAGGCGAGTGGCAT	3240
QY	3241	CAATTGCTGAGCTGTTTTATPAAGTTTCAAAATTCGAGCTTGTGTATAGGGAGAAAAA	3300
Db	3241	CAATTGCTGAGCTGTTTTATPAAGTTTCAAAATTCGAGCTTGTGTATAGGGAGAAAAA	3300
QY	3301	ACAATTTAGAAAATTTGTAAAGTTTGTGTCTGTAATGTTATGGTACCCAGAAATGTCG	3360
Db	3301	ACAATTTAGAAAATTTGTGTAAAGTTTGTGTCTGTAATGTTATGGTACCCAGAAATGTCG	3360
QY	3361	GACGAGGAATTGAACAAAGGACAGGTTTGATTGTTAAATGGCAAAAAAAAAAAAAA	3420
Db	3361	GACGAGGAATTGAACAAAGGACAGGTTTGATTGTTAAATGGCAAAAAAAAAAAAAA	3420
QY	3421	AAAAAAAAAAAAAAAAAAAAA 3443	
Db	3421	AAAAAAAAAAAAAAAAAAAAA 3443	

RESULT 2
US-10-209-059-29
; Sequence 29, Application US/10209059
; Publication NO. US2003016383A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.

; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; FILE REFERENCE: Thereof
; CURRENT APPLICATION NUMBER: 0864R2
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 3028
; TYPE: DNA
; ORGANISM: Zea mays
US-10-209-059-29

Query Match 87.9%; Score 3028; DB 16; Length 3028;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3028; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 416 CACGAGTTCAACATCGACGACGAGATCAGCAGGCGAGCTGGAGGGCAACATGCGAAC 475
Db 1 CACGAGTTCAACATCGACGACGAGATCAGCAGGCGAGCTGGAGGGCAACATGCGAAC 60
QY 476 AGCCAGATCACGAGGGGATGCTGCACGGCAGATGAGCTACGGAGGGGCCCGCAGCAG 535
Db 61 AGCCAGATCACGAGGGGATGCTGCACGGCAGATGAGCTACGGAGGGGCCCGCAGCAG 120
QY 536 GCGCAGCGCAACACACCCCGCAGATCCCGCCCATCATCACCGGCTCCCGCTCCGTCGG 595
Db 121 GCGCAGCGCAACACACCCCGCAGATCCCGCCCATCATCACCGGCTCCCGCTCCGTCGG 180
QY 596 GTGAGCGGTGAGTTTCGATTACCAACCGGTATGGCCACGGCAGGTCTCGTTCCCTG 655
Db 181 GTGAGCGGTGAGTTTCGATTACCAACCGGTATGGCCACGGCAGGTCTCGTTCCCTG 240
QY 656 CACAAGCGCATCCATCGTACCTGTCTGACGCCAGGAGTGCCCAAGTGGGACGAGAG 715
Db 241 CACAAGCGCATCCATCGTACCTGTCTGACGCCAGGAGTGCCCAAGTGGGACGAGAG 300
QY 716 AAAGAAGTGAAGTGAAGGAGAGATGGAACGACTGGAAGTCCAAGCAGGCGCATCTCCGC 775
Db 301 AAAGAAGTGAAGTGAAGGAGAGATGGAACGACTGGAAGTCCAAGCAGGCGCATCTCCGC 360
QY 776 GCGCGGCCCATCCCGAAGCATGGACGCCGACGCTGGCACTGAAACGAGGCGAGGCGAG 835
Db 361 GCGCGGCCCATCCCGAAGCATGGACGCCGACGCTGGCACTGAAACGAGGCGAGGCGAG 420
QY 836 CCCTGTGCGAGGAGGTGTCGATCGGTGAGCAAGGTGAACCCGTACCGGATGGTGATC 895
Db 421 CCCTGTGCGAGGAGGTGTCGATCGGTGAGCAAGGTGAACCCGTACCGGATGGTGATC 480
QY 896 GTGGTGGCTCTCGTTGTGCTCGCTTCTTCTCCGCTACCGGTATCTCTGCAACCCCGTCCCG 955
Db 481 GTGGTGGCTCTCGTTGTGCTCGCTTCTTCTCCGCTACCGGTATCTCTGCAACCCCGTCCCG 540
QY 956 GAGCCCATCGGGCTGTGGCTGTCTCCATCATCTGCGAGATCTGGTTGGCCATCTCTCTGG 1015
Db 541 GAGCCCATCGGGCTGTGGCTGTCTCCATCATCTGCGAGATCTGGTTGGCCATCTCTCTGG 600
QY 1016 ATCTCGACAGTTCCCAAGTGGTTCCCATCGACCGCAGACGTACCTCGACCGCCTC 1075
Db 601 ATCTCGACAGTTCCCAAGTGGTTCCCATCGACCGCAGACGTACCTCGACCGCCTC 660
QY 1076 TCCTCAGGTACGAGAGGAGGGAGCCGCTGCTGTGTCGGCGGTGGAACCTGTTCTG 1135
Db 661 TCCTCAGGTACGAGAGGAGGGAGCCGCTGCTGTGTCGGCGGTGGAACCTGTTCTG 720
QY 1136 AGCAGGTGACCCGCTCAAGGAGCCCGCTGGTGACCGCCCAACACCGTGTCTCTCATC 1195

Db 721 AGCAGGTGAGCCCGCTCAAGGAGCCCGCTGGTGACCCCAACACCGTGTCTCTCATC 780
QY 1196 CTCGCCGTAGACTACCCCGTGGACAAGGTCTCTGCTGCTCTCCGACGACGCGGTGCG 1255
Db 781 CTCGCCGTAGACTACCCCGTGGACAAGGTCTCTGCTGCTCTCCGACGACGCGGTGCG 840
QY 1256 ATGCTGACGTTTCGAGTTCGCTGTGCGAGACGGCGGAGTTTCGCGCGCAAGTGGTGCCTTC 1315
Db 841 ATGCTGACGTTTCGAGTTCGCTGTGCGAGACGGCGGAGTTTCGCGCGCAAGTGGTGCCTTC 900
QY 1316 TGCAGAAGTTTCGGCATCGAGCCCGCGCCCGCGAGTTTCTACTTCTCGCTCAAGTTCGAC 1375
Db 901 TGCAGAAGTTTCGGCATCGAGCCCGCGCCCGCGAGTTTCTACTTCTCGCTCAAGTTCGAC 960
QY 1376 TACCTCAAGGACAAAGGTGAGCCCACTTCTGTGCGAGAGCGCGCCCATGAAGAGAGAG 1435
Db 961 TACCTCAAGGACAAAGGTGAGCCCACTTCTGTGCGAGAGCGCGCCCATGAAGAGAGAG 1020
QY 1436 TATGAGGAGTTTCAAGGTCCGGATCAACGCGCTGGTGGCCAAAGGCCATGAAGTTCGCGCA 1495
Db 1021 TATGAGGAGTTTCAAGGTCCGGATCAACGCGCTGGTGGCCAAAGGCCATGAAGTTCGCGCA 1080
QY 1496 GAGGGGTGATCATGAAGACGGCAGCCCGTGGCCCGGGAAACAAACCCCGGACACCCCC 1555
Db 1081 GAGGGGTGATCATGAAGACGGCAGCCCGTGGCCCGGGAAACAAACCCCGGACACCCCC 1140
QY 1556 GGCATGATCACGAGTTCCTGGGCCACAGCGCGGCGCACACACCGAGGCGCAACGAGCTG 1615
Db 1141 GGCATGATCCAGGTGTTCTGGGCCACAGCGCGGCGCACACACCGAGGCGCAACGAGCTG 1200
QY 1616 CCCCGCTCGTGTAGTCTTCCGCTGAGAAAGCCCGGGATTCAGACCAACAAGAAAGGCC 1675
Db 1201 CCCCGCTCGTGTAGTCTTCCGCTGAGAAAGCCCGGGATTCAGACCAACAAGAAAGGCC 1260
QY 1676 GCGCCATGAAACGCTCTGATTCGCTTCCCGCGTGTGACCAACCGGCGCATTCATGCTC 1735
Db 1261 GCGCCATGAAACGCTCTGATTCGCTTCCCGCGTGTGACCAACCGGCGCATTCATGCTC 1320
QY 1736 AACTTGGACTGTGATCACTACATCAACACAGCAAGGCCATCCGGAGGCGCATGTGCTTC 1795
Db 1321 AACTTGGACTGTGATCACTACATCAACACAGCAAGGCCATCCGGAGGCGCATGTGCTTC 1380
QY 1796 CTCTAGGACCTTCAGGTCCGCGGAGAGGTCTGCTAGTTCAGTTCCTCCGAGAGTTCGAC 1855
Db 1381 CTCTAGGACCTTCAGGTCCGCGGAGAGGTCTGCTAGTTCAGTTCCTCCGAGAGTTCGAC 1440
QY 1856 GGCATCGACGTGACGACCGATACGCTAACAGGMAACACCGTCTTCTTCGACATCAACATG 1915
Db 1441 GGCATCGACGTGACGACCGATACGCTAACAGGMAACACCGTCTTCTTCGACATCAACATG 1500
QY 1916 AAGGGGCTGGAGCGCATCCAAAGCCCGGTGTACGTCCGACAGGGTGCCTGTTCGCGCGC 1975
Db 1501 AAGGGGCTGGAGCGCATCCAAAGCCCGGTGTACGTCCGACAGGGTGCCTGTTCGCGCGC 1560
QY 1976 CAGCGGCTCTACGCGTACAAACCTCCAAAGGACCCAAAGAGGCCAAAGATGCTGACCTGC 2035
Db 1561 CAGCGGCTCTACGCGTACAAACCTCCAAAGGACCCAAAGAGGCCAAAGATGCTGACCTGC 1620
QY 2036 GACTGTGCGCGTGTTCGGCCGCAAGAGCGGMAACACGCCAAGACCGGGTGCCTGGAG 2095
Db 1621 GACTGTGCGCGTGTTCGGCCGCAAGAGCGGMAACACGCCAAGACCGGGTGCCTGGAG 1680
QY 2096 GGCACCGCTGATATGGAGTAGTAGCGAAGAGGATGCTCATGTGCCACATGAAGACTTC 2155
Db 1681 GGCACCGCTGATATGGAGTAGTAGCGAAGAGGATGCTCATGTGCCACATGAAGACTTC 1740
QY 2156 GAGAAGCGTTCCGGCAGTCCCGCGGCTTCGTCAGCTCGACGCTGATGAGGAGGCGCGC 2215
Db 1741 GAGAAGCGTTCCGGCAGTCCCGCGGCTTCGTCAGCTCGACGCTGATGAGGAGGCGCGC 1800
QY 2216 GTCCCTCTCTGTCGAGCCCCCGCGCTCCTCAAGGAGGCCATCCATGTCTCATGCTGCTC 2275

Db 1801 GTCCCTCCTTCTCGAGCCCGCCGCTCTCTCAAGGAGGCCATCCATGTATCATAGCTGC 1860
Qy 2276 GGCTACGAGGACAAAGCCGACTGGGGCTGGAGCTGGGGTGAGTCTACGGGTGATCAGC 2335
Db 1861 GGCTACGAGGACAAAGCCGACTGGGGCTGGAGCTGGGGTGAGTCTACGGGTGATCAGC 1920
Qy 2336 GAGGACATCTGAGCGGGTTCAAGATGCACTGCCCGGGTGGCGCTCGGTGTACTGATG 2395
Db 1921 GAGGACATCTGAGCGGGTTCAAGATGCACTGCCCGGGTGGCGCTCGGTGTACTGATG 1980
Qy 2396 CCGAAGCGGGCGGCTTCAAGGGGTCCGCGCCGATCAATCTATCGGACCGCTCTCAACGAG 2455
Db 1981 CCGAAGCGGGCGGCTTCAAGGGGTCCGCGCCGATCAATCTATCGGACCGCTCTCAACGAG 2040
Qy 2456 GTGCTCCGGTGGGCGCTCGGTCCGCTCGAGATCTTCTTCAAGCGGACACGCCCTCTG 2515
Db 2041 GTGCTCCGGTGGGCGCTCGGTCCGCTCGAGATCTTCTTCAAGCGGACACGCCCTCTG 2100
Qy 2516 TAGGGCTACAAGACGGAACCTCAAGTGGCTGGAGCGCTTCGCTACATCAACACCA 2575
Db 2101 TAGGGCTACAAGACGGAACCTCAAGTGGCTGGAGCGCTTCGCTACATCAACACCA 2160
Qy 2576 ATCTACCCCTTCACTCGCTCCGCTGCTCGCTTACTGCACTTCCCGCGCTCTGCTC 2635
Db 2161 ATCTACCCCTTCACTCGCTCCGCTGCTCGCTTACTGCACTTCCCGCGCTCTGCTC 2220
Qy 2636 CTCACCGGCAAGTTTCATCATGCCGCTCGATTAGCACGTTTCGCGAGCTTCTTTCATCGCC 2695
Db 2221 CTCACCGGCAAGTTTCATCATGCCGCTCGATTAGCACGTTTCGCGAGCTTCTTTCATCGCC 2280
Qy 2696 CTCTTATGTTCATCTTCGGAACGGGATCTCTGAGATGCGGTGGAGCGGGGTGAGATC 2755
Db 2281 CTCTTATGTTCATCTTCGGAACGGGATCTCTGAGATGCGGTGGAGCGGGGTGAGATC 2340
Qy 2756 GAGGAGTGTGGAGGAACGAGCTTCTGGTTCATCGCGCGGTGTCGCGCATCTCTTC 2815
Db 2341 GAGGAGTGTGGAGGAACGAGCTTCTGGTTCATCGCGCGGTGTCGCGCATCTCTTC 2400
Qy 2816 GCGTCTGTCAGGGCTGCTCAAGTCTCTCGCGGGATCGACACCAACTTTCACCGTCACC 2875
Db 2401 GCGTCTGTCAGGGCTGCTCAAGTCTCTCGCGGGATCGACACCAACTTTCACCGTCACC 2460
Qy 2876 TCCAGGCGCACCGCGAGGAGGACGAGTTCGCGAGCTCTACGCTTCAAGTGAAC 2935
Db 2461 TCCAGGCGCACCGCGAGGAGGACGAGTTCGCGAGCTCTACGCTTCAAGTGAAC 2520
Qy 2936 ACGCTCTCATCCGCGCCACACGCTGCTCATCATTAAGTTCATCGCGGTGCTGGCGCGC 2995
Db 2521 ACGCTCTCATCCGCGCCACACGCTGCTCATTAAGTTCATCGCGGTGCTGGCGCGC 2580
Qy 2996 ATCTCCGACGCGCATCAACCGGTACAGTCTCTGGGGGCGCTCTTTCGGCAAGCTCTTC 3055
Db 2581 ATCTCCGACGCGCATCAACCGGTACAGTCTCTGGGGGCGCTCTTTCGGCAAGCTCTTC 2640
Qy 3056 TTGCGCTTCTGGGTCTATGCTCACTCTACCGGTCTCTCAAGGGCTCATGGGGCGCCAG 3115
Db 2641 TTGCGCTTCTGGGTCTATGCTCACTCTACCGGTCTCTCAAGGGCTCATGGGGCGCCAG 2700
Qy 3116 AACAGGACGCCACCGTGTGTGTCATCTGGTTCATCTGCTGGCGCTCCATCTTCTCCCTG 3175
Db 2701 AACAGGACGCCACCGTGTGTGTCATCTGGTTCATCTGCTGGCGCTCCATCTTCTCCCTG 2760
Qy 3176 CTCTGGGTCTAGGATCGAGCTTTCATCTGTCAGGACCAAGGCGCCGAGCTCAGGAGTGT 3235
Db 2761 CTCTGGGTCTAGGATCGAGCTTTCATCTGTCAGGACCAAGGCGCCGAGCTCAGGAGTGT 2820
Qy 3236 GGCATCAATTGCTGAGCTGTATTATTAAGGTTCAAAATTTCTGGAGCTTGTGATAGGAGA 3295
Db 2821 GGCATCAATTGCTGAGCTGTATTATTAAGGTTCAAAATTTCTGGAGCTTGTGATAGGAGA 2880
Qy 3296 AAAAAACAATTTAGAAATTTGTAAGGTTGTGTCTGTAAATTTATGTTATGTTACCAAGAT 3355
Db 2881 AAAAAACAATTTAGAAATTTGTAAGGTTGTGTCTGTAAATTTATGTTATGTTACCAAGAT 2940

Qy 3356 TGTCCGACGAGGAATTGAACAAAGGACAAGGTTTGTATTTGTAATGCAAAAAA 3415
Db 2941 TGTCCGACGAGGAATTGAACAAAGGACAAGGTTTGTATTTGTAATGCAAAAAA 3000
Qy 3416 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 3001 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3028

RESULT 3

US-10-425-114-833
; Sequence 833, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 833
; LENGTH: 2687
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700093978_FLI
US-10-425-114-833

Query Match 75.0%; Score 2583; DB 17; Length 2687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2633; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 773 GGCGGGCGGCCCATCCGAAAGACATGGAGCGCGAGTGGCACTGAACGACGAGCGGAGG 832
Db 1 GGCGGGCGGCCCATCCGAAAGACATGGAGCGCGAGTGGCACTGAACGACGAGCGGAGG 60
Qy 833 CAGCGCTGTGAGGAGAGGTTCGATCGCGTTCGAGCAAGGTGAACCGTACCGGATGGTG 892
Db 61 CAGCGCTGTGAGGAGAGGTTCGATCGCGTTCGAGCAAGGTGAACCGTACCGGATGGTG 120
Qy 893 ATCGTGGTGGTCTCGTGTGCTCGCTTCTTCTCGGTACCGTATCTCGACCCCGTC 952
Db 121 ATCGTGGTGGTCTCGTGTGCTCGCTTCTTCTCGGTACCGTATCTCGACCCCGTC 180
Qy 953 CCGGACGCCATCGGGCTGTGGCTCTCATCATCTCGAGATCTGGTTCGCCATCTCC 1012
Db 181 CCGGACGCCATCGGGCTGTGGCTCTCATCATCTCGAGATCTGGTTCGCCATCTCC 240
Qy 1013 TGGATCTCGACAGTTCCCAAGTGTTCCTCCATCGACCGGAGAGTACCTCGACCGC 1072
Db 241 TGGATCTCGACAGTTCCCAAGTGTTCCTCCATCGACCGGAGAGTACCTCGACCGC 300
Qy 1073 CTCTCCCTCAGGTACGAGAGGAGGAGCGCTCGTCTGCTCGGGTGGACCTGTTC 1132
Db 301 CTCTCCCTCAGGTACGAGAGGAGGAGCGCTCGTCTGCTCGGGTGGACCTGTTC 360
Qy 1133 GTGAGCAGGTGACCGCGCTCAAGGAGCGCGCTGGTGAACCGCAACCGGTCTCTCC 1192
Db 361 GTGAGCAGGTGACCGCGCTCAAGGAGCGCGCTGGTGAACCGCAACCGGTCTCTCC 420
Qy 1193 ATCTCGCGGTAGACTACCGCGTGGAGAGGTCTCTGTGTACGTCTCCGACGCGCGC 1252
Db 421 ATCTCGCGGTAGACTACCGCGTGGAGAGGTCTCTGTGTACGTCTCCGACGCGCGC 480
Qy 1253 TCGATCTGACGTTGAGTCTGTCGAGACCGCGAGTTCGCGCGCAAGTGGGTGCCCC 1312

Db	481	TCGATGCTGACGTTTCGAGTGCCTGTCGGAGACGGCCGAGTTTCGGCGCCAAAGTGGGTGCC	540
Qy	1313	TTCTGCAAGAAGTTTCGGCATCGAGCCCGCGCCCCGGAGTTCTACTTCTCGCTCAAGGTC	1372
Db	541	TTCTGCAAGAAGTTTCGGCATCGAGCCCGCGCCCCGGAGTTCTACTTCTCGCTCAAGGTC	600
Qy	1373	GACTACCTCAAGGACAAGGTGACGCCACCTTCTGTGAGAGAGCCCGGCATGAAGAGA	1432
Db	601	GACTACCTCAAGGACAAGGTGACGCCACCTTCTGTGAGAGAGCCCGGCATGAAGAGA	660
Qy	1433	GAGTATCAGGAGTTCAAGGTCGGATCAACGCGCTGTGTGCGCAAGGCCATGAAGGTGCCG	1492
Db	661	GAGTATCAGGAGTTCAAGGTCGGATCAACGCGCTGTGTGCGCAAGGCCATGAAGGTGCCG	720
Qy	1493	GCAGAGGGTGGATCATGAAGGACGGCAACGCGCTGTGCGCGGGAACAACACCCGCGACAC	1552
Db	721	GCAGAGGGTGGATCATGAAGGACGGCAACGCGCTGTGCGCGGGAACAACACCCGCGACAC	780
Qy	1553	CCCGGCATGATTCAGAGTTCTCTGGGCCACAGCGGGCGGCCACGACACCGAGGGCAACAG	1612
Db	781	CCCGGCATGATTCAGAGTTCTCTGGGCCACAGCGGGCGGCCACGACACCGAGGGCAACAG	840
Qy	1613	CTGCGCCGCTCGTGTAAGTCTCCGCTGAGAGCGCCCGGGATTCGAGCACCAAGAAG	1672
Db	841	CTGCGCCGCTCGTGTAAGTCTCCGCTGAGAGCGCCCGGGATTCGAGCACCAAGAAG	900
Qy	1673	GCCGGGCCATGAACGCTCTGATTCGCGTCTTCGCGCTGTGTGACCAACGCGCCATTCATG	1732
Db	901	GCCGGGCCATGAACGCTCTGATTCGCGTCTTCGCGCTGTGTGACCAACGCGCCATTCATG	960
Qy	1733	CTCAACTTGGACTGTGTATCACTACATCAACACAGCAAGGCCATTCGCGGAGGCCATGTGC	1792
Db	961	CTCAACTTGGACTGTGTATCACTACATCAACACAGCAAGGCCATTCGCGGAGGCCATGTGC	1020
Qy	1793	TTCTCTATGACCCCTCAGGTCGGCGGAAAGTCTGTACGTTTCAAGTTCCTCGCAGAGGTTTC	1852
Db	1021	TTCTCTATGACCCCTCAGGTCGGCGGAAAGTCTGTACGTTTCAAGTTCCTCGCAGAGGTTTC	1080
Qy	1853	GACGGCATCGACGTGACGACCGATACGCTAAACAGGAACACCGTCTCTTCGACATCAAC	1912
Db	1081	GACGGCATCGACGTGACGACCGATACGCTAAACAGGAACACCGTCTCTTCGACATCAAC	1140
Qy	1913	ATGAAGGGGTGGACGGCATCCAAAGGCCCGGTGTACGTCGGGACAGGGTCGCTGTTCCGG	1972
Db	1141	ATGAAGGGGTGGACGGCATCCAAAGGCCCGGTGTACGTCGGGACAGGGTCGCTGTTCCGG	1200
Qy	1973	GGCAGGCGCTCTACGGCTACAACCCCTCCAAAGGACCCAAAGAGGCCAAGATGTTGAC	2032
Db	1201	GGCAGGCGCTCTACGGCTACAACCCCTCCAAAGGACCCAAAGAGGCCAAGATGTTGAC	1260
Qy	2033	TGCGACTGTGCCGCTCTTCGGCCGCAAGAAGCGGAAACACGCCAAGGACGGGTCGCCG	2092
Db	1261	TGCGACTGTGCCGCTCTTCGGCCGCAAGAAGCGGAAACACGCCAAGGACGGGTCGCCG	1320
Qy	2093	GAGGCAACCGCTGATATGGGAGTAGATAGCGCAAGGAGATGCTCATGTCCCACTGAAC	2152
Db	1321	GAGGCAACCGCTGATATGGGAGTAGATAGCGCAAGGAGATGCTCATGTCCCACTGAAC	1380
Qy	2153	TTGAGAAAGCGGTTTCGGGCAAGTCCGCGCGGTTTGTCACTCGACGCTGATGAGGAGAGC	2212
Db	1381	TTGAGAAAGCGGTTTCGGGCAAGTCCGCGCGGTTTGTCACTCGACGCTGATGAGGAGAGC	1440
Qy	2213	GGCGTCCCTCTTGTGAGCCCGCGCGCTCTCTAAGGAGGCCATCCATGTTCATCAGC	2272
Db	1441	GGCGTCCCTCTTGTGAGCCCGCGCGCTCTCTAAGGAGGCCATCCATGTTCATCAGC	1500
Qy	2273	TGCGGCTACAGGACAAGAACCGACTGGGGGTGGAGCTGGGGTGGATCTACGGGTGCATC	2332
Db	1501	TGCGGCTACAGGACAAGAACCGACTGGGGGTGGAGCTGGGGTGGATCTACGGGTGCATC	1560
Qy	2333	ACGAGGACATCTCTGAACGGGTTCAAGATGACATGCGCGGGTGGCGCTCCGTGTACTGC	2392
Db	1561	ACGAGGACATCTCTGAACGGGTTCAAGATGACATGCGCGGGTGGCGCTCCGTGTACTGC	1620

DEC 11 1964

RESULTS 4
US-10-425-115-139789

03-10-423-113-139789 ; Sequence 139789, Application US/10425115

```
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; NUMBER OF SEQ ID NOS: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 139789
; LENGTH: 2830
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_58974C.1
US-10-425-115-139789

Query Match      5.0%; Score 173; DB 18; Length 2830;
Best Local Similarity 98.9%; Pred. No. 7.7e-77;
Matches 373; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1457 ATCAACGCGCTGGTGGCCAAAGGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1516
Db 826 ATCAACGCGCTGGTGGCCAAAGGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 885

Qy 1517 GGACAGCGCTGGTGGCCAAAGGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1576
Db 886 GGACAGCGCTGGTGGCCAAAGGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 945

Qy 1577 GGCCACAGCGCGCGCCAGCACCAACCGCGACCAACCGCGAGGCGCCGCTGCTGTTCTCC 1636
Db 946 GGCCACAGCGCGCGCCAGCACCAACCGCGAGGCGCCGCTGCTGTTCTCC 1005

Qy 1637 CGTGAGAGCGCGCGGATTCACAGCACCAAGAGGCGCGCGGCGCCGCTGCTGATT 1696
Db 1006 CGTGAGAGCGCGCGGATTCACAGCACCAAGAGGCGCGCGGCGCCGCTGCTGATT 1065

Qy 1697 CGCGTCTCCGCGTGTGACCAACCGCGGATTCATGCTCACTTTGGACTGTGATCACTAC 1756
Db 1066 CGCGTCTCCGCGTGTGACCAACCGCGGATTCATGCTCACTTTGGACTGTGATCACTAC 1125

Qy 1757 ATCAACAACAGCAAGGCCATCCGGAGGCCATGTCTTCTCATGGACCCCTCAGGTCCGC 1816
Db 1126 ATCAACAACAGCAAGGCCATCCGGAGGCCATGTCTTCTCATGGACCCCTCAGGTCCGC 1185

Qy 1817 CGGAAGGTCTGCTAGT 1833
Db 1186 CGGAAGGTCTGCTAGT 1202

RESULT 5
US-10-425-114-24930
; Sequence 24930, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 24930
; LENGTH: 3589
; TYPE: DNA
```

```
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3689-221-C10_FLI
US-10-425-114-24930

Query Match      5.0%; Score 173; DB 17; Length 3589;
Best Local Similarity 98.9%; Pred. No. 7.5e-77;
Matches 373; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1457 ATCAACGCGCTGGTGGCCAAAGGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1516
Db 1615 ATCAACGCGCTGGTGGCCAAAGGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1674

Qy 1517 GGACAGCGCTGGTGGCCAAAGGCCATGAAGTGGCGGAGGGGTGGATCATGAAGTGGTCTG 1576
Db 1675 GGACAGCGCTGGTGGCCAAAGGCCATGAAGTGGCGGAGGGGTGGATCATGAAGTGGTCTG 1734

Qy 1577 GGCCACAGCGCGCGCCAGCACCAACCGCGACCAACCGCGAGGCGCCGCTGCTGTTCTCC 1636
Db 1735 GGCCACAGCGCGCGCCAGCACCAACCGAGGGGACAGGCTGCCGCCCTGCTGTTCTCC 1794

Qy 1637 CGTGAGAGCGCGCGGATTCACAGCACCAAGAGGCGCGGCGCCATGAACGCTTCTGATT 1696
Db 1795 CGTGAGAGCGCGCGGATTCACAGCACCAAGAGGCGCGGCGCCATGAACGCTTCTGATT 1854

Qy 1697 CGCGTCTCCGCGTGTGACCAACCGCGGATTCATGCTCACTTTGGACTGTGATCACTAC 1756
Db 1855 CGCGTCTCCGCGTGTGACCAACCGCGGATTCATGCTCACTTTGGACTGTGATCACTAC 1914

Qy 1757 ATCAACAACAGCAAGGCCATCCGGAGGCCATGTCTTCTCATGGACCCCTCAGGTCCGC 1816
Db 1915 ATCAACAACAGCAAGGCCATCCGGAGGCCATGTCTTCTCATGGACCCCTCAGGTCCGC 1974

Qy 1817 CGGAAGGTCTGCTAGT 1833
Db 1975 CGGAAGGTCTGCTAGT 1991

RESULT 6
US-10-425-115-179459
; Sequence 179459, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 179459
; LENGTH: 335
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(335)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_9524C.1
US-10-425-115-179459

Query Match      3.6%; Score 124; DB 18; Length 335;
Best Local Similarity 100.0%; Pred. No. 6.7e-52;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCGTCGCCCTCGCTCGCAATCGCAATCTGTGAGCACCTGAGGGGTGGAGGCCGAG 60
Db 126 CTGCGTCGCCCTCGCTCGCAATCGCAATCTGTGAGCACCTGAGGGGTGGAGGCCGAG 185
```



```
QY 61 AGCTAGCTAGCAGCGCGCTCGCGCGCGATGAGAGCCAGCGCGCGTGTGTCGCGG 120
Db 186 AGCTAGCTAGCAGCGCGCTCGCGCGCGATGAGAGCCAGCGCGCGTGTGTCGCGG 245
QY 121 CTCG 124
Db 246 CTCG 249

RESULT 7
US-10-767-701-10062
; Sequence 10062, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 10062
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(999)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS50500_1
US-10-767-701-10062

Query Match 3.4%; Score 116; DB 18; Length 999;
Best Local Similarity 100.0%; Pred. No. 6.8e-48;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2312 GCGTGGATCTACGGGTGATCAGGAGGACATCTGACGGGGTTCAAGATGCACTGCCG 2371
Db 35 GCGTGGATCTACGGGTGATCAGGAGGACATCTGACGGGGTTCAAGATGCACTGCCG 94

QY 2372 GCGTGGCGTCCGTTACTGCATCCGAAACGGCGCGCGTTCAAGGGTGGCGCC 2427
Db 95 GCGTGGCGTCCGTTACTGCATCCGAAACGGCGCGCGTTCAAGGGTGGCGCC 150

RESULT 8
US-10-425-115-10141
; Sequence 10141, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 10141
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_10924C.1
US-10-425-115-10141

Query Match 3.2%; Score 110; DB 18; Length 633;
Best Local Similarity 99.4%; Pred. No. 7.9e-45;
```

```
Matches 160; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 281 TGCCGCGCTGTGTAGAGTACGAGCGCCGCGAGGGCACGACAGACTGCCCGCCAGTCAAG 340
Db 436 TGCCGCGCTGTGTAGAGTACGAGCGCCGCGAGGGCACGACAGACTGCCCGCCAGTCAAG 495
QY 341 ACGCGCTACAAGCGCTCAAGGGGAGCCGAGGGTTGCCGGGAGCATCACGAGGAGGAC 400
Db 496 ACGCGCTACAAGCGCTCAAGGGGAGCCGAGGGTCCGCGGAGCATCACGAGGAGGAC 555
QY 401 ATCGACGACCTGAGCAGCAGTTCACATCGACGAGGAA 441
Db 556 ATCGACGACCTGAGCAGCAGTTCACATCGACGAGGAA 596

RESULT 9
US-10-437-963-15093/c
; Sequence 15093, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 15093
; LENGTH: 3448
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_20969C.1
US-10-437-963-15093

Query Match 2.4%; Score 83; DB 18; Length 3448;
Best Local Similarity 100.0%; Pred. No. 3.3e-31;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 813 CACTGAACGACGAGCGGAGCGCGCTGTGCGAGGAAGGTGCGATCGCGTCGAGCAAG 872
Db 2646 CACTGAACGACGAGCGGAGCGCGCTGTGCGAGGAAGGTGCGATCGCGTCGAGCAAG 2587

QY 873 TGAACCCGTACCGATGGTGATC 895
Db 2586 TGAACCCGTACCGATGGTGATC 2564

RESULT 10
US-10-767-701-29317
; Sequence 29317, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 29317
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Sorghum bicolor
```


; FEATURE:
; OTHER INFORMATION: Clone ID: 9298823
US-10-767-701-29317

Query Match 1.7%; Score 59; DB 18; Length 549;
Best Local Similarity 100.0%; Pred. No. 5.8e-19; Indels 0; Gaps 0;
Matches 59; Conservative 0; Mismatches 0;

Qy 1733 CTCACCTTGACTGTGATCACTACATCAACACAGCAAGGCCATCCGGAGGCCATGTG 1791
Db 1 CTCACCTTGACTGTGATCACTACATCAACACAGCAAGGCCATCCGGAGGCCATGTG 59

RESULT 11

US-10-425-115-139788
; Sequence 139788, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 139788
; LENGTH: 488
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_58973C.1
US-10-425-115-139788

Query Match 1.7%; Score 58; DB 18; Length 488;
Best Local Similarity 100.0%; Pred. No. 1.9e-18; Indels 0; Gaps 0;
Matches 58; Conservative 0; Mismatches 0;

Qy 1077 CCCTCAGTACGAGAGGAGGAGCGCGTGTGCTGTGCGCGGTGGACCTGTTGCT 1134
Db 201 CCCTCAGTACGAGAGGAGGAGCGCGTGTGCTGTGCGCGGTGGACCTGTTGCT 258

RESULT 12

US-09-900-237-5
; Sequence 5, Application US/09900237
; Patent No. US2002012024A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: Bb1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1189
; TYPE: DNA
; ORGANISM: Zea mays
US-09-900-237-5

Query Match 1.5%; Score 50; DB 9; Length 1189;
Best Local Similarity 100.0%; Pred. No. 2e-14; Indels 0; Gaps 0;
Matches 50; Conservative 0; Mismatches 0;

Qy 2771 AACGAGCAGTTCTGGGTTCATCGCGCGGTGTCCGGCGATCTCTTCGCCGT 2820
Db 485 AACGAGCAGTTCTGGGTTCATCGCGCGGTGTCCGGCGATCTCTTCGCCGT 534

RESULT 13

US-10-425-114-2510
; Sequence 2510, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 2510
; LENGTH: 2351
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700214676_FLI
US-10-425-114-2510

Query Match 1.5%; Score 50; DB 17; Length 2351;
Best Local Similarity 100.0%; Pred. No. 1.9e-14; Indels 0; Gaps 0;
Matches 50; Conservative 0; Mismatches 0;

Qy 2771 AACGAGCAGTTCTGGGTTCATCGCGCGGTGTCCGGCGATCTCTTCGCCGT 2820
Db 1674 AACGAGCAGTTCTGGGTTCATCGCGCGGTGTCCGGCGATCTCTTCGCCGT 1723

RESULT 14

US-10-425-115-137279
; Sequence 137279, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 137279
; LENGTH: 2684
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2684)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_56676C.1
US-10-425-115-137279

Query Match 1.5%; Score 50; DB 18; Length 2684;
Best Local Similarity 100.0%; Pred. No. 1.8e-14; Indels 0; Gaps 0;
Matches 50; Conservative 0; Mismatches 0;

Qy 2771 AACGAGCAGTTCTGGGTTCATCGCGCGGTGTCCGGCGATCTCTTCGCCGT 2820
Db 1843 AACGAGCAGTTCTGGGTTCATCGCGCGGTGTCCGGCGATCTCTTCGCCGT 1892

```
RESULT 15
US-10-209-059-25
; Sequence 25, Application US/10209059
; Publication No. US20030163838A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; PRIOR FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 3470
; TYPE: DNA
; ORGANISM: Zea mays
US-10-209-059-25

Query Match          1.5%; Score 50; DB 16; Length 3470;
Best Local Similarity 100.0%; Pred. No. 1.8e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2771 AACGAGCAGTTCGGGTCTCGCGCGGTGTCGCGCATCTCTTCGCCCGT 2820
    |||||||
Db 2786 AACGAGCAGTTCGGGTCTCGCGCGGTGTCGCGCATCTCTTCGCCCGT 2835

RESULT 16
US-10-627-132-25
; Sequence 25, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 3470
; TYPE: DNA
; ORGANISM: Zea mays
US-10-627-132-25

Query Match          1.5%; Score 50; DB 17; Length 3470;
Best Local Similarity 100.0%; Pred. No. 1.8e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2771 AACGAGCAGTTCGGGTCTCGCGCGGTGTCGCGCATCTCTTCGCCCGT 2820
    |||||||
Db 2786 AACGAGCAGTTCGGGTCTCGCGCGGTGTCGCGCATCTCTTCGCCCGT 2835
```

```
RESULT 17
US-09-955-999-23
; Sequence 23, Application US/09955999
; Publication No. US20030036505A1
; GENERAL INFORMATION:
; APPLICANT: Barash et al.
; TITLE OF INVENTION: Signal Transduction Pathway Component Polynucleotides, Polypeptides, Antibodies, and Methods Based Thereon
; FILE REFERENCE: PT086PI
; CURRENT APPLICATION NUMBER: US/09/955,999
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 907
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (625)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (787)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-955-999-23

Query Match          1.4%; Score 48; DB 10; Length 907;
Best Local Similarity 100.0%; Pred. No. 2.1e-13;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3396 TAAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
    |||||||
Db 859 TAAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 906

RESULT 18
US-10-437-963-39652
; Sequence 39652, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 39652
; LENGTH: 1657
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_43170C.1
US-10-437-963-39652

Query Match          1.4%; Score 48; DB 18; Length 1657;
Best Local Similarity 100.0%; Pred. No. 2e-13;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3396 TAAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
```


; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 378
; LENGTH: 1826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1799)..(1799)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-378

Query Match 1.4%; Score 47; DB 15; Length 1826;
Best Local Similarity 100.0%; Pred. No. 6.3e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3397 AATGGCAA 3443
|||||
Db 1740 AATGGCAA 1786

RESULT 24
US-10-425-115-55122/c
; Sequence 55122, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: Plants
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 55122
; LENGTH: 586
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_150269C.1
US-10-425-115-55122

Query Match 1.3%; Score 46; DB 18; Length 586;
Best Local Similarity 100.0%; Pred. No. 2.3e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3398 AATGGCAA 3443
|||||
Db 84 AATGGCAA 39

RESULT 25
US-09-925-298-401
; Sequence 401, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 401
; LENGTH: 189
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (11)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (162)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (165)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (166)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (187)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-298-401

Query Match 1.3%; Score 45; DB 9; Length 189;
Best Local Similarity 100.0%; Pred. No. 8.1e-12;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAA 3443
|||||
Db 104 ATGGCAA 148

RESULT 26
US-10-102-806-401
; Sequence 401, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103Pic1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 401
; LENGTH: 189
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (11)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (162)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (165)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature

```
; LOCATION: (166)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (187)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-102-806-401

Query Match
Best Local Similarity 1.3%; Score 45; DB 14; Length 189;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 104 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 148

RESULT 27
US-10-425-115-102183
; Sequence 102183, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 102183
; LENGTH: 356
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(358)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_24699C.1
US-10-425-115-102183

Query Match
Best Local Similarity 1.3%; Score 45; DB 18; Length 358;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 91 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 135

RESULT 28
US-09-918-995-11420
; Sequence 11420, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11420
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

; LOCATION: (1)...(480)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-11420

Query Match
Best Local Similarity 1.3%; Score 45; DB 10; Length 480;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 208 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 252

RESULT 29
US-10-425-115-52426
; Sequence 52426, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 52426
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_147812C.1
US-10-425-115-52426

Query Match
Best Local Similarity 1.3%; Score 45; DB 18; Length 711;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 77 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 121

RESULT 30
US-10-425-115-105831
; Sequence 105831, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 105831
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(864)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_28018C.1
US-10-425-115-105831

Query Match
Best Local Similarity 1.3%; Score 45; DB 18; Length 864;
```

Search completed: March 12, 2005, 10:42:43
Job time : 1226 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2005, 02:20:28 ; Search time 7010 Seconds
(without alignments)
18695.475 Million cell updates/sec

Title: US-10-627-132-29

Perfect score: 3443

Sequence: 1 ctgcgtgcctgcctgcga.....aaaaaaaaaaaaaaaa 3443

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 25

Total number of hits satisfying chosen parameters: 596789

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

1: gb_est1.*

2: gb_est2.*

3: gb_est3.*

4: gb_est4.*

5: gb_est5.*

6: gb_est6.*

7: gb_gss1.*

8: gb_gss2.*

9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	745	21.6	943	9	CG454544
C 2	694	20.2	872	8	CG427726
C 3	643	18.7	949	9	CG211396
C 4	583	16.9	822	9	CG198455
C 5	368	10.7	480	2	BE050991
C 6	356	10.3	631	8	CG427725
C 7	356	10.3	739	9	CG145359
C 8	356	10.3	782	9	CG371975
C 9	356	10.3	828	9	CG619840
C 10	356	10.3	835	9	CG300337
C 11	356	10.3	867	9	CG081564
C 12	356	10.3	917	9	CG263400
C 13	356	10.3	945	9	CG300346
C 14	349	10.1	818	8	CG446974
C 15	332	9.6	452	9	CG656950
C 16	320	9.3	608	6	CA183623
C 17	320	9.3	624	6	CA173049
C 18	318	9.2	808	9	CG436514
C 19	305	8.9	748	9	CG454501
C 20	269	7.8	640	9	CG705263
C 21	269	7.8	833	8	CG446976
C 22	269	7.8	841	8	BZ733711
C 23	266	7.7	651	8	CC389096
C 24	266	7.7	815	9	CC705271

CC655942	818	7.7	266	C 25	CG655942
CG343883	7.7	266	266	C 26	CG343883
CG145361	7.7	261	261	C 27	CG145361
BE050992	7.7	246	246	C 28	BE050992
BZ733703	821	236	236	C 29	BZ733703
CG371985	758	214	214	C 30	CG371985
CG430615	841	214	214	C 31	CG430615
CG263391	900	214	214	C 32	CG263391
CG343868	919	214	214	C 33	CG343868
CG434216	936	214	214	C 34	CG434216
CG434218	943	214	214	C 35	CG434218
CG685594	636	206	206	C 36	CG685594
CG685588	641	206	206	C 37	CG685588
CG619834	702	206	206	C 38	CG619834
CG081561	811	206	206	C 39	CG081561
CC000780	869	203	203	C 40	CC000780
CA173106	543	200	200	C 41	CA173106
CC374709	349	183	183	C 42	CC374709
CA194818	628	182	182	C 43	CA194818
CG712515	280	178	178	C 44	CG712515
CG081196	418	173	173	C 45	CG081196
BZ640404	592	173	173	C 46	BZ640404
CG651626	844	173	173	C 47	CG651626
CG615542	904	173	173	C 48	CG615542
CG370129	999	173	173	C 49	CG370129
CF433508	727	167	167	C 50	CF433508
CA165585	589	166	166	C 51	CA165585
CA155363	706	158	158	C 52	CA155363
SCAQR2310	710	149	149	C 53	SCAQR2310
SCQGR2304	702	148	148	C 54	SCQGR2304
BZ640411	736	146	146	C 55	BZ640411
CG381495	868	146	146	C 56	CG381495
CG606677	907	146	146	C 57	CG606677
CG381504	631	134	134	C 58	CG381504
CA133799	834	134	134	C 59	CA133799
CO519653	354	131	131	C 60	CO519653
CC374710	653	126	126	C 61	CC374710
CA235459	227	125	125	C 62	CA235459
CG712518	687	122	122	C 63	CG712518
CA127439	744	122	122	C 64	CA127439
CA208554	781	122	122	C 65	CA208554
CA129877	484	118	118	C 66	CA129877
CF650564	552	116	116	C 67	CF650564
BE357401	789	116	116	C 68	BE357401
CF430024	578	113	113	C 69	CF430024
CA268135	693	113	113	C 70	CA268135
SCCCFL412	693	113	113	C 71	SCCCFL412
CL161384	412	108	108	C 72	CL161384
CL222797	505	107	107	C 73	CL222797
BI141114	533	107	107	C 74	BI141114
BE357339	777	107	107	C 75	BE357339
CG081198	675	107	107	C 76	CG081198
CL167974	892	102	102	C 77	CL167974
CA253951	935	102	102	C 78	CA253951
SCRLFL410	959	101	101	C 79	SCRLFL410
CG359736	288	101	101	C 80	CG359736
CG038978	288	98	98	C 81	CG038978
AQ844522	179	93	93	C 82	AQ844522
BE125071	600	92	92	C 83	BE125071
CA172230	715	92	92	C 84	CA172230
CA244716	755	92	92	C 85	CA244716
CA069753	902	92	92	C 86	CA069753
CG722849	902	92	92	C 87	CG722849
CG198038	935	92	92	C 88	CG198038
CG615547	982	92	92	C 89	CG615547
OGWBL547V	733	92	92	C 90	OGWBL547V
CC704091	958	85	85	C 91	CC704091
CC350870	747	84	84	C 92	CC350870
CG519652	727	83	83	C 93	CG519652
CO519652	3168	83	83	C 94	CO519652
CB646549	943	83	83	C 95	CB646549
CL976367	566	77	77	C 96	CL976367
CG660675	74	74	74	C 97	CG660675
OGUL007TH	585	4	4	C 98	OGUL007TH
SCSFSR107	585	4	4	C 99	SCSFSR107
BU255931	585	4	4	C 100	BU255931

98 72 2.1 981 9 CG370118 CG370118 OGYCQ52TH
99 71 2.1 409 9 CL161383 CL161383
100 71 2.1 429 6 CA291358 CA291358 SCCCN200

ALIGNMENTS

RESULT 1
CG454544/c 943 bp DNA linear GSS 17-SEP-2003
LOCUS PUIKA96TDB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTA0601023,
DEFINITION genomic survey sequence.

ACCESSION CG454544
VERSION CG454544.1 GI:34839544
KEYWORDS GSS.

SOURCE

ORGANISM

Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 943)

REFERENCE

AUTHORS

Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.

Maize Genomics Consortium

Unpublished (2003)

Other GSSs: PUIKA96TDB PUIKA96TBB

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

FEATURES

source

1..943
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTA0601023"
/clone_lib="ZM_0.6_1.0_KB"
/notes="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

ORIGIN

Query Match 21.6%; Score 745; DB 9; Length 943;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 745; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2662 GATTAGCAGTTCCGCGAGCTCTTCTTCATCGCCCTCTTCATGTCCTTCGCGAGCGG 2721
|||||

Db 903 GATTAGCAGTTCCGCGAGCTCTTCTTCATCGCCCTCTTCATGTCCTTCGCGAGCGG 844
|||||

QY 2722 CATCTCGAGATCGGTGGAGCGGTGAGCATCGAGGAGTGTGGAGGAACGAGCAGTT 2781
|||||

Db 843 CATCTCGAGATCGGTGGAGCGGTGAGCATCGAGGAGTGTGGAGGAACGAGCAGTT 784
|||||

QY 2782 CTGGGTTCATCGCGCGGTGTCGCGCATCTTTGCGCGTGTGTCAGGGCTGCTCAAGT 2841
|||||

Db 783 CTGGGTTCATCGCGCGGTGTCGCGCATCTTTGCGCGTGTGTCAGGGCTGCTCAAGT 724
|||||

QY 2842 CCTCGCGGATCGACCAACTTCCCGTCACTCCAGGCCACCGGCGAGGAGCA 2901
|||||

Db 723 CCTCGCGGATCGACCAACTTCCCGTCACTCCAGGCCACCGGCGAGGAGCA 664
|||||

QY 2902 CGAGTTTCGCGAGCTCTAGCCCTTCAAGTGGACCGCTCTTCATCCCGCCACCAACGCT 2961
|||||

Db 663 CGAGTTTCGCGAGCTCTAGCCCTTCAAGTGGACCGCTCTTCATCCCGCCACCAACGCT 604
|||||

QY 2962 GCTCATCTTAACGTATCGGGCTGTGGCGGCGCATCTCGAGCCCATCAACACGGGTA 3021
|||||

Db 603 GCTCATCTTAACGTATCGGGCTGTGGCGGCGCATCTCGAGCCCATCAACACGGGTA 544
|||||

QY 3022 CAGTCCCTGGGGGCCCTCTTCGGCAAGCTCTTCTTCGCTTCTGGGTTCATCGTCCACCT 3081
|||||

Db 543 CAGTCCCTGGGGGCCCTCTTCGGCAAGCTCTTCTTCGCTTCTGGGTTCATCGTCCACCT 484
|||||

QY 3082 CTACCCGTTCTCAAGGGGCTCATGGGGGCCAGAACAGAGCGCCACCGTTGTTGTCAT 3141
|||||

Db 483 CTACCCGTTCTCAAGGGGCTCATGGGGGCCAGAACAGAGCGCCACCGTTGTTGTCAT 424
|||||

QY 3142 CTGGTCCATCTTCGTGGCTTCATCTTCTTCCTGCTCTGGGTTCAGATCGACCTTTTCAT 3201
|||||

Db 423 CTGGTCCATCTTCGTGGCTTCATCTTCTTCCTGCTCTGGGTTCAGATCGACCTTTTCAT 364
|||||

QY 3202 CGTCAGGACCAAGGGGCCGAGCTCAGGAGTGTGCATCTTCTGAGCTGTTTATTA 3261
|||||

Db 363 CGTCAGGACCAAGGGGCCGAGCTCAGGAGTGTGCATCTTCTGAGCTGTTTATTA 304
|||||

QY 3262 AGGTTCAAAATTCGAGCTTGTGCATAGGAGAAAAAACAATTTAGAAAATTTGTAAG 3321
|||||

Db 303 AGGTTCAAAATTCGAGCTTGTGCATAGGAGAAAAAACAATTTAGAAAATTTGTAAG 244
|||||

QY 3322 GTTGTGTCCTGTAATGTTATGTTACCCAGAAATTCGAGAGGAATTTGAACAAAGGA 3381
|||||

Db 243 GTTGTGTCCTGTAATGTTATGTTACCCAGAAATTCGAGAGGAATTTGAACAAAGGA 184
|||||

QY 3382 CAAGTTTGTATTTAAATGGCAAA 3406
|||||

Db 183 CAAGTTTGTATTTAAATGGCAAA 159
|||||

RESULT 2

CC427726/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

1..872

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBTA407B09"

/clone_lib="ZM_0.6_1.0_KB"

/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

ORIGIN

Query Match

20.2%; Score 694; DB 8; Length 872;

VERSION	CG198455.1	GI:34089516
KEYWORDS	GSS.	
SOURCE	Zea mays	
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.	
REFERENCE	1 (bases 1 to 822)	
AUTHORS	Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennetzen, J.	
TITLE	Maize Genomics Consortium	
JOURNAL	Unpublished (2003)	
COMMENT	Contact: Cathy Whitelaw	
FEATURES	TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@tigr.org Seq primer: TF Class: sheared ends. Location/Qualifiers 1..822 /organism="Zea mays" /mol_type="genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZMBETA0601023" /clone_lib="ZM 0.6 1.0 KB" /note="Vector: pCRI-TOPO; Site 1: EcoRI; 0.6-1.0 kb high Cot selected genomic DNA library"	
ORIGIN	Query Match 16.9%; Score 583; DB 9; Length 822; Best Local Similarity 99.8%; Pred. No. 8.7e-263; Matches 633; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	2773	CGAGCAGTTCGGGTGATCGCGCGGTGTCGCGCATCTCTTCGCGTGTGAGGCGCT 2832
DB	793	CGAGCAGTTCGGGTGATCGCGCGGTGTCGCGCATCTCTTCGCGTGTGAGGCGCT 734
QY	2833	GCTCAAGGTCCTCGCGGATCGACCACTTCAAGTCACTTCAAGGCGGCGGCGCA 2892
DB	733	GCTCAAGGTCCTCGCGGATCGACCACTTCAAGTCACTTCAAGGCGGCGGCGCA 674
QY	2893	CGAGGACGAGTTCGCGAGCTCTACGCTTCAAGTGGACCACTCTCTATCCCGCC 2952
DB	673	CGAGGACGAGTTCGCGAGCTCTACGCTTCAAGTGGACCACTCTCTATCCCGCC 614
QY	2953	CACACGCTGCTCATATTAACGTATCGGCGTGTGGCGGATCTCGACGCCATCAA 3012
DB	613	CACACGCTGCTCATATTAACGTATCGGCGTGTGGCGGATCTCGACGCCATCAA 554
QY	3013	CAACGGGTACAGTCTCGGGGCCCTCTTCGCAAGCTTCTTCGCTTCTGGGTGAT 3072
DB	553	CAACGGGTACAGTCTCGGGGCCCTCTTCGCAAGCTTCTTCGCTTCTGGGTGAT 494
QY	3073	CGTCCACCTCTACCCGTTCTCAAGGGGCTTCATGGGGCGCAGAACAGGACGCC 3132
DB	493	CGTCCACCTCTACCCGTTCTCAAGGGGCTTCATGGGGCGCAGAACAGGACGCC 434
QY	3133	TGTTGTGATGTCATTTCTGTCGCTTCATCTCTTCGCTTCTTCGCTTCTGGGTG 3192
DB	433	TGTTGTGATGTCATTTCTGTCGCTTCATCTCTTCGCTTCTTCGCTTCTGGGTG 374
QY	3193	CCCTTTTCATGTCAGGACCAAGGCGCGGACGTTCAGGAGTGGGATCAATTTGCTG 3252
DB	373	CCCTTTTCATGTCAGGACCAAGGCGCGGACGTTCAGGAGTGGGATCAATTTGCTG 314
QY	3253	TGTTTATTAAGGTTCAAAATTCGAGCTTGTGCATAGGAGAGAGAGAGAGAGAG 3312
DB	313	TGTTTATTAAGGTTCAAAATTCGAGCTTGTGCATAGGAGAGAGAGAGAGAGAG 254

QY	3313	TTTTGTAGGTTGTTGTTCTGTAATTTATGATCCAGAAATTTGTCGACGAGGAATTG 3372
DB	253	TTTTGTAGGTTGTTGTTCTGTAATTTATGATCCAGAAATTTGTCGACGAGGAATTG 194
QY	3373	AACAAGGACAAAGGTTTGAATTTAAATGGCAAA 3406
DB	193	AACAAGGACAAAGGTTTGAATTTAAATGGCAAA 160
RESULT 5	BE050991	480 bp mRNA linear EST 18-MAR-2002
LOCUS	za71d11.b50	Maize Glume cDNAs Library Zea mays cDNA clone za71d11
DEFINITION	5', mRNA sequence.	
ACCESSION	BE050991	
VERSION	BE050991.1	GI:8368046
KEYWORDS	EST.	
SOURCE	Zea mays	
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.	
REFERENCE	1 (bases 1 to 480)	
AUTHORS	O'Shaughnessy, A.L., Habermann, K., de la Bastide, M., Huang, B.N., Nascimento, L.U., Schutz, K., Matero, A., Swaby, I., See, L.-H., Preston, R.R., Rodriguez, M.A., Shah, R.S., Shekher, M., Spiegel, L.A., Vil, M.D., Dedhia, N.N. and McCombie, W.R.	
TITLE	Expressed sequence tags from Zea mays (maize)	
JOURNAL	Unpublished (2000)	
COMMENT	Contact: W. Richard McCombie Lita Annenberg Hazen Genome Sequencing Center Cold Spring Harbor Laboratory PO Box 100, Cold Spring Harbor, NY 11724, USA Tel: 516 367 8884 Fax: 516 367 8874 Email: mcombie@cshl.org Plate: za71 row: d column: 11 Seq primer: -40M13ForUniv High quality sequence stop: 480. Location/Qualifiers 1..480 /organism="Zea mays" /mol_type="mRNA" /db_xref="taxon:4577" /clone="za71d11" /clone_lib="Maize Glume cDNAs Library" /note="Vector: Lambda Zap II (Stratagene); Site 1: XhoI; Site 2: EcoRI; Resistance: Ampicillin; Autoexcision: pBluescript SK (+/-); Titer: 7 x 10e-9 pfu/mL (as of 9/28/94); Library source: John Doebley, B73"	
FEATURES	Query Match 10.7%; Score 368; DB 2; Length 480; Best Local Similarity 99.6%; Pred. No. 1.5e-161; Matches 468; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
ORIGIN	1594 CGACACCGAGGCAACGAGCTGCCCGCTCGTGACGTCTCCGTCGAGAGCGCCGG 1653	
QY	1594	CGACACCGAGGCAACGAGCTGCCCGCTCGTGACGTCTCCGTCGAGAGCGCCGG 70
DB	11	CGACACCGAGGCAACGAGCTGCCCGCTCGTGACGTCTCCGTCGAGAGCGCCGG 70
QY	1654	ATTCCAGCACCAAGAGGCGCGCCATGACGCTCTGATTCGGCTCTCCGCGTCT 1713
DB	71	ATTCCAGCACCAAGAGGCGCGCCATGACGCTCTGATTCGGCTCTCCGCGTCT 130
QY	1714	GACCAACGCGCATTCATGCTCAACTTGGACTGTGATCACTACTACATCAACAAGCAAGGC 1773
DB	131	GACCAACGCGCATTCATGCTCAACTTGGACTGTGATCACTACTACATCAACAAGCAAGGC 190
QY	1774	CATCCGGGAGGCATGTGCTTCTCATGACCCCTCAGGTTCGGCCGGAAGGTTCTGTACGT 1833
DB	191	CATCCGGGAGGCATGTGCTTCTCATGACCCCTCAGGTTCGGCCGGAAGGTTCTGTACGT 250
QY	1834	TCAGTTCCCGCAGAGGTTTCGAGCGGATCGACGTGCACGACCGCATCGCTAACAGGAACAC 1893

```

Db      251 TCAGTTCGCGAGAGTTGACGGCATCGAGTGCACGACGATACGTAACAGGAACAC 310
Qy      1894 CGTCTTCTTGACATCAACATGAAGGGCTGGAGGCATCAAGGCCGGGTGTACGTCGG 1953
Db      311 CGTCTTCTTGACATCAACATGAAGGGCTGGAGGCATCAAGGCCGGGTGTACGTCGG 370
Qy      1954 GACAGGGTGGCTGTTCCGGGCGCAGGGCTTACGGCTACAACTCCCAAGGACCCAA 2013
Db      371 GACAGGGTGGCTGTTCCGGGCGCAGGGCTTACGGCTACAACTCCCAAGGACCCAA 430
Qy      2014 GAGGCCCAAGATGCTGCACTGCTGCGCTGCTGCGCGCGCAAGA 2063
Db      431 GAGGCCCAAGATGCTGCACTGCTGCGCTGCTGCGCGCGCAAGA 480

RESULT 6
CC427725
LOCUS   CC427725
DEFINITION PUHAM29TB ZM 0.6_1.0 KB Zea mays genomic clone ZMMBta407E09,
          genomic survey sequence.
ACCESSION CC427725
VERSION   CC427725.1 GI:30907815
KEYWORDS  GSS.
SOURCE    Zea mays
ORGANISM  Zea mays
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
          clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 631)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
          Reenick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
          Bennetzen,J.
TITLE     Maize Genomics Consortium
JOURNAL   Unpublished (2003)
COMMENT   Other GSSs: PUHAM29TD
          Contact: Cathy Whitelaw
          TIGR
          9712 Medical Center Drive, Rockville, MD 20850, USA
          Tel: 301-838-5843
          Fax: 301-838-0208
          Email: whitelaw@tigr.org
          Seq primer: TR
          Class: sheared ends.
FEATURES  Location/Qualifiers
            source
              1..631
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone_lib="ZMMBta407E09"
                /note="Vector: PCR4-TOFO; Site 1: EcoRI; 0.6-1.0 kb high
                CoT selected genomic DNA library"

ORIGIN
Query Match 10.3%; Score 356; DB 8; Length 631;
Best Local Similarity 100.0%; Pred. No. 6.5e-156;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2307 AGCTGGGGTGGATCTACGGGTGATCAGGAGGATCCTGACGGGGTTCAAGATGCAC 2366
Db      54 AGCTGGGGTGGATCTACGGGTGATCAGGAGGATCCTGACGGGGTTCAAGATGCAC 113
Qy      2367 GCCCGGGTGGCGCTCCGTGTACTGTCATGCCAGAGCGGGCGCTTCAAGGGGTGCGCGC 2426
Db      114 GCCCGGGTGGCGCTCCGTGTACTGTCATGCCAGAGCGGGCGCTTCAAGGGGTGCGCGC 173
Qy      2427 CGATCAATCTATCGGACCGTCTCAACAGAGTGTCTCCGGTGGCGCTCGGTTCGAGA 2486
Db      174 CGATCAATCTATCGGACCGTCTCAACAGAGTGTCTCCGGTGGCGCTCGGTTCGAGA 233
Qy      2487 TCCTCTTCAGCGGCACAGCCCCCTGCTGTACGGCTACAGAGACGGCAACCTCAAGTGGC 2546

```

```

Db      234 TCCTTTTACCGCGCACAGCCCCCTGCTGTACGGCTACAAGAACGGCAACCTCAAGTGGC 293
Qy      2547 TGGAGCGCTTTCGCTTACATCAAGCACCAACCATCTACCCCTTACCTCGCTCCCGCTGCTCG 2606
Db      294 TGGAGCGCTTTCGCTTACATCAAGCACCAACCATCTACCCCTTACCTCGCTCCCGCTGCTCG 353
Qy      2607 CCTACTGCACCTTCCCGCGCTGCTGCTCTCACCAGGCAAGTTTCATCATGCGCTCG 2662
Db      354 CCTACTGCACCTTCCCGCGCTGCTGCTCTCACCAGGCAAGTTTCATCATGCGCTCG 409

RESULT 7
CG145359
LOCUS   CG145359
DEFINITION PUJEW43TB ZM 0.6_1.0 KB Zea mays genomic clone ZMMBta0652H13,
          genomic survey sequence.
ACCESSION CG145359
VERSION   CG145359.1 GI:34036142
KEYWORDS  GSS.
SOURCE    Zea mays
ORGANISM  Zea mays
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
          clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 739)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
          Reenick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
          Bennetzen,J.
TITLE     Maize Genomics Consortium
JOURNAL   Unpublished (2003)
COMMENT   Other GSSs: PUJEW43TD
          Contact: Cathy Whitelaw
          TIGR
          9712 Medical Center Drive, Rockville, MD 20850, USA
          Tel: 301-838-5843
          Fax: 301-838-0208
          Email: whitelaw@tigr.org
          Seq primer: TR
          Class: sheared ends.
FEATURES  Location/Qualifiers
            source
              1..739
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone_lib="ZMMBta0652H13"
                /clone_lib="ZM 0.6_1.0 KB"
                /note="Vector: PCR4-TOFO; Site 1: EcoRI; 0.6-1.0 kb high
                CoT selected genomic DNA library"

ORIGIN
Query Match 10.3%; Score 356; DB 9; Length 739;
Best Local Similarity 100.0%; Pred. No. 6.4e-156;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2307 AGCTGGGGTGGATCTACGGGTGATCAGGAGGATCCTGACGGGGTTCAAGATGCAC 2366
Db      112 AGCTGGGGTGGATCTACGGGTGATCAGGAGGATCCTGACGGGGTTCAAGATGCAC 171
Qy      2367 GCCCGGGTGGCGCTCCGTGTACTGTCATGCCAGAGCGGGCGCTTCAAGGGGTGCGCGC 2426
Db      172 GCCCGGGTGGCGCTCCGTGTACTGTCATGCCAGAGCGGGCGCTTCAAGGGGTGCGCGC 231
Qy      2427 CGATCAATCTATCGGACCGTCTCAACAGAGTGTCTCCGGTGGCGCTCGGTTCGAGA 2486
Db      232 CGATCAATCTATCGGACCGTCTCAACAGAGTGTCTCCGGTGGCGCTCGGTTCGAGA 291
Qy      2487 TCCTTTTACCGCGGCACAGCCCCCTGCTGTACGGCTACAAGAACGGCAACCTCAAGTGGC 2546
Db      292 TCCTTTTACCGCGGCACAGCCCCCTGCTGTACGGCTACAAGAACGGCAACCTCAAGTGGC 351
Qy      2547 TGGAGCGCTTTCGCTTACATCAACACCAACATCTACCCCTTTCACCTCGCTCCCGCTGCTCG 2606

```

```
||||| 352 TGGAGCGCTTGGCTTACATCAACACCACCATCTACCCCTTACCTCGCTCCGCGTCTCG 411
|||||
QY 2607 CCTACTGCACCTCCCGCGCTGCTCTCTACCGGCAAGTTTCATCATGCGGTGC 2662
|||||
Db 412 CCTACTGCACCTCCCGCGCTGCTCTCTACCGGCAAGTTTCATCATGCGGTGC 467
|||||

RESULT 8
CG371975/c
LOCUS
DEFINITION
CG371975
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 782)
AUTHORS
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
TITLE
Unpublished (2002)
JOURNAL
COMMENT
Other GSSs: OGIAD62TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
Location/Qualifiers
source
1..782
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMBMa0715L04"
/clone_lib="ZM.0.7.1.5 KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 10.3%; Score 356; DB 9; Length 782;
Best Local Similarity 100.0%; Pred. No. 6.4e-156;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2307 AGCTGGGGTGGATCTACGGGTGCATCAGGAGACATCTCTGACGGGGTTCAAGATGCACT 2366
|||||
Db 609 AGCTGGGGTGGATCTACGGGTGCATCAGGAGACATCTCTGACGGGGTTCAAGATGCACT 550
|||||
QY 2367 GCCCGGGTGGCGCTCGGTGATCTGTCATGCCGAAGCGGGCGGTTCAAGGGGTCCGCGC 2426
|||||
Db 549 GCCCGGGTGGCGCTCGGTGATCTGTCATGCCGAAGCGGGCGGTTCAAGGGGTCCGCGC 490
|||||
QY 2427 CGATCAATCTATCGGACCGGTCTCAACACAGGTGCTCCGGTGGCGCTCGGTCCGTCGAGA 2486
|||||
Db 489 CGATCAATCTATCGGACCGGTCTCAACACAGGTGCTCCGGTGGCGCTCGGTCCGTCGAGA 430
|||||
QY 2487 TCTTCTTACCGCGCACAGCCCCCTGCTGTCACGGCTACAGAAGCGGCAACCTCAAGTGGC 2546
|||||
Db 429 TCTTCTTACCGCGCACAGCCCCCTGCTGTCACGGCTACAGAAGCGGCAACCTCAAGTGGC 370
|||||
QY 2547 TGGAGCGCTTCGGCTTACATCAACACCATCTACCCCTTACCTCGCTCCCGCTGCTCG 2606
|||||
Db 369 TGGAGCGCTTCGGCTTACATCAACACCATCTACCCCTTACCTCGCTCCCGCTGCTCG 310
|||||
QY 2607 CCTACTGCACCTCCCGCGCTGCTGCTCTCTACCGGCAAGTTTCATCATGCGGTGC 2662
|||||
```

```
||||| 309 CCTACTGCACCTCCCGCGCTGCTCTCTACCGGCAAGTTTCATCATGCGGTGC 254
|||||

RESULT 9
CG619840/c
LOCUS
DEFINITION
CG619840
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 828)
AUTHORS
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
TITLE
Unpublished (2002)
JOURNAL
COMMENT
Other GSSs: OGULU56TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
Location/Qualifiers
source
1..828
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMBMa0470J15"
/clone_lib="ZM.0.7.1.5 KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 10.3%; Score 356; DB 9; Length 828;
Best Local Similarity 100.0%; Pred. No. 6.4e-156;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2307 AGCTGGGGTGGATCTACGGGTGCATCAGGAGACATCTCTGACGGGGTTCAAGATGCACT 2366
|||||
Db 453 AGCTGGGGTGGATCTACGGGTGCATCAGGAGACATCTCTGACGGGGTTCAAGATGCACT 394
|||||
QY 2367 GCCCGGGTGGCGCTCCGTTGTCATGCCGAAGCGGGCGGTTCAAGGGGTCCGCGC 2426
|||||
Db 393 GCCCGGGTGGCGCTCCGTTGTCATGCCGAAGCGGGCGGTTCAAGGGGTCCGCGC 334
|||||
QY 2427 CGATCAATCTATCGGACCGGTCTCAACACAGGTGCTCCGGTGGCGCTCGGTCCGTCGAGA 2486
|||||
Db 333 CGATCAATCTATCGGACCGGTCTCAACACAGGTGCTCCGGTGGCGCTCGGTCCGTCGAGA 274
|||||
QY 2487 TCTTCTTACCGCGCACAGCCCCCTGCTGTCACGGCTACAGAAGCGGCAACCTCAAGTGGC 2546
|||||
Db 273 TCTTCTTACCGCGCACAGCCCCCTGCTGTCACGGCTACAGAAGCGGCAACCTCAAGTGGC 214
|||||
QY 2547 TGGAGCGCTTCGGCTTACATCAACACCATCTACCCCTTACCTCGCTCCCGCTGCTCG 2606
|||||
Db 213 TGGAGCGCTTCGGCTTACATCAACACCATCTACCCCTTACCTCGCTCCCGCTGCTCG 154
|||||
QY 2607 CCTACTGCACCTCCCGCGCTGCTGCTCTCTACCGGCAAGTTTCATCATGCGGTGC 2662
|||||
Db 153 CCTACTGCACCTCCCGCGCTGCTGCTCTCTACCGGCAAGTTTCATCATGCGGTGC 98
|||||
```

```
RESULT 10
CG0300337
LOCUS
DEFINITION
  CG2AK93TH ZM 0.7_1.5_KB Zea mays genomic clone ZMMBma0743P17,
  genomic survey sequence.
ACCESSION
  CG300337
VERSION
  CG300337.1 GI:34214551
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
  1 (bases 1 to 835)
AUTHORS
  Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
  Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
  Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  Consortium for Maize Genomics
TITLE
  Unpublished (2002)
JOURNAL
  Unpublished
COMMENT
  Other_GSSs: CG2AK93TV
  Contact: Cathy Whitelaw
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TR
  Class: sheared ends.
FEATURES
  source
    1..835
    Location/Qualifiers
      /organism="Zea mays"
      /mol_type="genomic DNA"
      /strain="B73"
      /db_xref="taxon:4577"
      /clone="ZMMBma0743P17"
      /notes="Vector: pBCSk-; Site 1: HincII; 0.7-1.5 kb
      methylation filtered genomic DNA library"
ORIGIN
  Query Match          10.3%; Score 356; DB 9; Length 835;
  Best Local Similarity 100.0%; Pred. No. 6.4e-156;
  Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 2307 AGCTGGGGTGGATCTACGGGTGCGATCAGGAGGACATCTCTGACGGGGTTCAAGATGCAC 2366
  DB 373 AGCTGGGGTGGATCTACGGGTGCGATCAGGAGGACATCTCTGACGGGGTTCAAGATGCAC 432
  QY 2367 GCCCGGGGTGGCGTCCGCTGCTACTGCGATGCGAGCGGGCGGCTTCAAGGGGTCCGCGC 2426
  DB 433 GCCCGGGGTGGCGTCCGCTGCTACTGCGATGCGAGCGGGCGGCTTCAAGGGGTCCGCGC 492
  QY 2427 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCCGGTGGCGCTGGGGTCCGTCGAGA 2486
  DB 493 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCCGGTGGCGCTGGGGTCCGTCGAGA 552
  QY 2487 TCTTCTTTCAGCCGACAGCCCGCTGCTGCTACCGGTCTACGAGACGGCAACCTCAAGTGGC 2546
  DB 553 TCTTCTTTCAGCCGACAGCCCGCTGCTGCTACCGGTCTACGAGACGGCAACCTCAAGTGGC 612
  QY 2547 TGGAGCGCTTGGCTATACACACACCATCTACCCCTTACCTGCTCCCGTGGCTCG 2606
  DB 613 TGGAGCGCTTGGCTATACACACACCATCTACCCCTTACCTGCTCCCGTGGCTCG 672
  QY 2607 CCTACTGCACCTCCCGCGCTGCTGCTCTCCTACCGGCAAGTTTCATATGCCGTCG 2662
  DB 673 CCTACTGCACCTCCCGCGCTGCTGCTCTCCTACCGGCAAGTTTCATATGCCGTCG 728
  RESULT 11
  CG081564/c
  LOCUS
  DEFINITION
    CG2AK93TH ZM 0.7_1.5_KB Zea mays genomic clone ZMMBma0743P17,
    genomic survey sequence.
  ACCESSION
    CG300337
  VERSION
    CG300337.1 GI:34214551
  KEYWORDS
    GSS.
  SOURCE
    Zea mays
    ORGANISM
      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
      clade; Panicoideae; Andropogoneae; Zea.
  REFERENCE
    1 (bases 1 to 835)
  AUTHORS
    Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
    Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
    Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
    Consortium for Maize Genomics
  TITLE
    Unpublished (2002)
  JOURNAL
    Unpublished
  COMMENT
    Other_GSSs: CG2AK93TV
    Contact: Cathy Whitelaw
    TIGR
    9712 Medical Center Drive, Rockville, MD 20850, USA
    Tel: 301-838-5843
    Fax: 301-838-0208
    Email: whitelaw@tigr.org
    Seq primer: TR
    Class: sheared ends.
  FEATURES
    source
      1..835
      Location/Qualifiers
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /strain="B73"
        /db_xref="taxon:4577"
        /clone="ZMMBma0743P17"
        /notes="Vector: pBCSk-; Site 1: HincII; 0.7-1.5 kb
        methylation filtered genomic DNA library"
  ORIGIN
    Query Match          10.3%; Score 356; DB 9; Length 835;
    Best Local Similarity 100.0%; Pred. No. 6.4e-156;
    Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY 2307 AGCTGGGGTGGATCTACGGGTGCGATCAGGAGGACATCTCTGACGGGGTTCAAGATGCAC 2366
    DB 373 AGCTGGGGTGGATCTACGGGTGCGATCAGGAGGACATCTCTGACGGGGTTCAAGATGCAC 432
    QY 2367 GCCCGGGGTGGCGTCCGCTGCTACTGCGATGCGAGCGGGCGGCTTCAAGGGGTCCGCGC 2426
    DB 433 GCCCGGGGTGGCGTCCGCTGCTACTGCGATGCGAGCGGGCGGCTTCAAGGGGTCCGCGC 492
    QY 2427 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCCGGTGGCGCTGGGGTCCGTCGAGA 2486
    DB 493 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCCGGTGGCGCTGGGGTCCGTCGAGA 552
    QY 2487 TCTTCTTTCAGCCGACAGCCCGCTGCTGCTACCGGTCTACGAGACGGCAACCTCAAGTGGC 2546
    DB 553 TCTTCTTTCAGCCGACAGCCCGCTGCTGCTACCGGTCTACGAGACGGCAACCTCAAGTGGC 612
    QY 2547 TGGAGCGCTTGGCTATACACACACCATCTACCCCTTACCTGCTCCCGTGGCTCG 2606
    DB 613 TGGAGCGCTTGGCTATACACACACCATCTACCCCTTACCTGCTCCCGTGGCTCG 672
    QY 2607 CCTACTGCACCTCCCGCGCTGCTGCTCTCCTACCGGCAAGTTTCATATGCCGTCG 2662
    DB 673 CCTACTGCACCTCCCGCGCTGCTGCTCTCCTACCGGCAAGTTTCATATGCCGTCG 728
  RESULT 12
  CG263400/c
  LOCUS
  DEFINITION
    CG263400 ZM 0.7_1.5_KB Zea mays genomic clone ZMMBma0663N07,
    genomic survey sequence.
  ACCESSION
    CG263400
  VERSION
    CG263400.1 GI:34175541
  KEYWORDS
    GSS.
```

```
genomic survey sequence.
CG081564
CG081564.1 GI:33963858
GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
  1 (bases 1 to 867)
AUTHORS
  Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
  Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
  Bennetzen,J.
  Maize Genomics Consortium
  Unpublished (2003)
  Other_GSSs: PUJFC40TB
  Contact: Cathy Whitelaw
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TF
  Class: sheared ends.
FEATURES
  source
    1..867
    Location/Qualifiers
      /organism="Zea mays"
      /mol_type="genomic DNA"
      /strain="B73"
      /db_xref="taxon:4577"
      /clone="ZMMBma0654G07"
      /notes="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
      COT selected genomic DNA library"
ORIGIN
  Query Match          10.3%; Score 356; DB 9; Length 867;
  Best Local Similarity 100.0%; Pred. No. 6.4e-156;
  Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 2307 AGCTGGGGTGGATCTACGGGTGCGATCAGGAGGACATCTCTGACGGGGTTCAAGATGCAC 2366
  DB 411 AGCTGGGGTGGATCTACGGGTGCGATCAGGAGGACATCTCTGACGGGGTTCAAGATGCAC 352
  QY 2367 GCCCGGGGTGGCGTCCGCTGCTACTGCGATGCGAGCGGGCGGCTTCAAGGGGTCCGCGC 2426
  DB 351 GCCCGGGGTGGCGTCCGCTGCTACTGCGATGCGAGCGGGCGGCTTCAAGGGGTCCGCGC 292
  QY 2427 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCCGGTGGCGCTGGGGTCCGTCGAGA 2486
  DB 291 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCCGGTGGCGCTGGGGTCCGTCGAGA 232
  QY 2487 TCTTCTTTCAGCCGACAGCCCGCTGCTGCTACCGGTCTACGAGACGGCAACCTCAAGTGGC 2546
  DB 231 TCTTCTTTCAGCCGACAGCCCGCTGCTGCTACCGGTCTACGAGACGGCAACCTCAAGTGGC 172
  QY 2547 TGGAGCGCTTGGCTATACACACACCATCTACCCCTTACCTGCTCCCGTGGCTCG 2606
  DB 171 TGGAGCGCTTGGCTATACACACACCATCTACCCCTTACCTGCTCCCGTGGCTCG 112
  QY 2607 CCTACTGCACCTCCCGCGCTGCTGCTCTCCTACCGGCAAGTTTCATATGCCGTCG 2662
  DB 111 CCTACTGCACCTCCCGCGCTGCTGCTCTCCTACCGGCAAGTTTCATATGCCGTCG 56
  RESULT 12
  CG263400/c
  LOCUS
  DEFINITION
    CG263400 ZM 0.7_1.5_KB Zea mays genomic clone ZMMBma0663N07,
    genomic survey sequence.
  ACCESSION
    CG263400
  VERSION
    CG263400.1 GI:34175541
  KEYWORDS
    GSS.
```

SOURCE	Zea mays	clade; Panicoidae; Andropogoneae; Zea.
ORGANISM	Zea mays	1 (bases 1 to 945)
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD	Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
AUTHORS	1 (bases 1 to 917)	Consortium for Maize Genomics
TITLE	Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.	Unpublished (2002)
JOURNAL	Consortium for Maize Genomics	Other GSSs: OG2AK93TH
COMMENT	Unpublished (2002)	Contact: Cathy Whitelaw
OTHER GSSs:	OGXDU76TH	TIGR
CONTACT:	Cathy Whitelaw	9712 Medical Center Drive, Rockville, MD 20850, USA
TELE:	301-838-5843	Tel: 301-838-5843
FAX:	301-838-0208	Fax: 301-838-0208
EMAIL:	whitelaw@tigr.org	Email: whitelaw@tigr.org
SEQ PRIMER:	TF	Seq primer: TF
CLASS:	sheared ends.	Class: sheared ends.
FEATURES	Location/Qualifiers	Location/Qualifiers
SOURCE	1..917	1..945
	/organism="Zea mays"	/organism="Zea mays"
	/mol_type="genomic DNA"	/mol_type="genomic DNA"
	/strain="B73"	/strain="B73"
	/db_xref="taxon:4577"	/db_xref="taxon:4577"
	/clone_lib="ZMMBma0663N07"	/clone_lib="ZMMBma0743P17"
	/clone_lib="ZM 0.7 1.5 KB"	/clone_lib="ZM 0.7 1.5 KB"
	/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"	/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN		
Query Match	10.3%; Score 356; DB 9; Length 917;	10.3%; Score 356; DB 9; Length 945;
Best Local Similarity	100.0%; Pred. No. 6.3e-156;	100.0%; Pred. No. 6.3e-156;
Matches	356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	2307 AGCTGGGGTGATCTACGGGTGATCAGGAGGACATCTCTGACGGGGTTCAAGATGCAT 2366	2307 AGCTGGGGTGATCTACGGGTGATCAGGAGGACATCTCTGACGGGGTTCAAGATGCAT 2366
DB	571 AGCTGGGGTGATCTACGGGTGATCAGGAGGACATCTCTGACGGGGTTCAAGATGCAT 512	820 AGCTGGGGTGATCTACGGGTGATCAGGAGGACATCTCTGACGGGGTTCAAGATGCAT 761
QY	2367 GCGCGGGTGGCGTCCGGTACTGATCCGAAGCGGGCGGTTCAAGGGGTGGCGC 2426	2367 GCGCGGGTGGCGTCCGGTACTGATCCGAAGCGGGCGGTTCAAGGGGTGGCGC 2426
DB	511 GCGCGGGTGGCGTCCGGTACTGATCCGAAGCGGGCGGTTCAAGGGGTGGCGC 452	760 GCGCGGGTGGCGTCCGGTACTGATCCGAAGCGGGCGGTTCAAGGGGTGGCGC 701
QY	2427 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCCGGTGGGGTCCGTCGAGA 2486	2427 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCCGGTGGGGTCCGTCGAGA 2486
DB	451 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCCGGTGGGGTCCGTCGAGA 392	700 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCCGGTGGGGTCCGTCGAGA 641
QY	2487 TCTTCTTACCGGGCAGCGCCCTGCTGTAAGGCTACAAGAACGGCAACCTCAAGTGGC 2546	2487 TCTTCTTACCGGGCAGCGCCCTGCTGTAAGGCTACAAGAACGGCAACCTCAAGTGGC 2546
DB	391 TCTTCTTACCGGGCAGCGCCCTGCTGTAAGGCTACAAGAACGGCAACCTCAAGTGGC 332	640 TCTTCTTACCGGGCAGCGCCCTGCTGTAAGGCTACAAGAACGGCAACCTCAAGTGGC 581
QY	2547 TGGAGGGCTTCGGCTACATCAACACCAACATCTACCCCTTACCTCGCTCCGCTGCTG 2606	2547 TGGAGGGCTTCGGCTACATCAACACCAACATCTACCCCTTACCTCGCTCCGCTGCTG 2606
DB	331 TGGAGGGCTTCGGCTACATCAACACCAACATCTACCCCTTACCTCGCTCCGCTGCTG 272	580 TGGAGGGCTTCGGCTACATCAACACCAACATCTACCCCTTACCTCGCTCCGCTGCTG 521
QY	2607 CCTACTGCACCTCCCGCGGTGCTGCTCTCTCACCGGCAAGTTTCATCATCGCGTCG 2662	2607 CCTACTGCACCTCCCGCGGTGCTGCTCTCTCACCGGCAAGTTTCATCATCGCGTCG 2662
DB	271 CCTACTGCACCTCCCGCGGTGCTGCTCTCTCACCGGCAAGTTTCATCATCGCGTCG 216	520 CCTACTGCACCTCCCGCGGTGCTGCTCTCTCACCGGCAAGTTTCATCATCGCGTCG 465
RESULT 13		
CG300346/c		
LOCUS	CG300346 TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBma0743P17,	CC446974
DEFINITION	genomic survey sequence.	FUHLI43TB ZM_0.6_1.0 KB Zea mays genomic clone ZMMBta477H13,
ACCESSION	CG300346	CC446974
VERSION	CG300346.1 GI:34214560	CC446974.1 GI:30954665
KEYWORDS	GSS.	GSS
SOURCE	Zea mays	Zea mays
ORGANISM	Zea mays	Zea mays
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
	1 (bases 1 to 818)	1 (bases 1 to 818)
REFERENCE	Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and	Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
AUTHORS		

Bennetzen, J.
Maize Genomics Consortium
Unpublished (2003)
Other_GSSs: PUHLI437D
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

Location/Qualifiers
source
1..818
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBTa477H13"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

FEATURES
source
1..818
Location/Qualifiers

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBTa477H13"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

Query Match 10.1%; Score 349; DB 8; Length 818;
Best Local Similarity 100.0%; Pred. No. 1.3e-152;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

Qy 1081 CAGGTACGAGAGGAGGAGCGTCTGCTGCGCGGTGGACCTGTTCTGAGGAC 1140
Db 357 CAGGTACGAGAGGAGGAGCGTCTGCTGCTGCGCGGTGGACCTGTTCTGAGGAC 298
Qy 1141 GGTGGACCCGCTCAAGGAGCGCGGTGGTGAACCGCAACCGTCTCTCCATCCTCGC 1200
Db 297 GGTGGACCCGCTCAAGGAGCGCGGTGGTGAACCGCAACCGTCTCTCCATCCTCGC 238
Qy 1201 CGTAGACTACCCGTGACAAAGTCTCTGCTAGTCTCCGACGCGCGGTGATGCT 1260
Db 237 CGTAGACTACCCGTGACAAAGTCTCTGCTAGTCTCCGACGCGCGGTGATGCT 178
Qy 1261 GACGTTTCGAGTCGCTGTCGAGAGCGCGCGGTGTCGCGCAAGTGGTGGCCCTTCGCA 1320
Db 177 GACGTTTCGAGTCGCTGTCGAGAGCGCGCGGTGTCGCGCAAGTGGTGGCCCTTCGCA 118
Qy 1321 GAAGTTCGGCATCGAGCCCGCGCGCGCGGTGTCGCTTCGCTCAAGGTCGACTACCT 1380
Db 117 GAAGTTCGGCATCGAGCCCGCGCGCGCGGTGTCGCTTCGCTCAAGGTCGACTACCT 58
Qy 1381 CAAGGACAAAGTGCAGCCACCTTCTGTCAGGAGCGCGCGCATGAAG 1429
Db 57 CAAGGACAAAGTGCAGCCACCTTCTGTCAGGAGCGCGCGCATGAAG 9

RESULT 15
CC656950
LOCUS
DEFINITION
OGLAD37TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0307H02,
genomic survey sequence.

ACCESSION
CC656950
VERSION
CC656950.1 GI:32060243
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays

REFERENCE
AUTHORS
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Frazer, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nurnberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGLAD37TH

Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.

Location/Qualifiers
source
1..452
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBMA0307H02"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES

Location/Qualifiers
source
1..452
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBMA0307H02"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 9.6%; Score 332; DB 9; Length 452;
Best Local Similarity 100.0%; Pred. No. 1.3e-144;
Matches 332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1081 CAGGTACGAGAGGAGGAGCGTCTGCTGCGCGGTGGACCTGTTCTGAGGAC 1140
Db 121 CAGGTACGAGAGGAGGAGCGTCTGCTGCTGCGCGGTGGACCTGTTCTGAGGAC 180
Qy 1141 GGTGGACCCGCTCAAGGAGCGCGGTGGTGAACCGCAACCGTCTCTCCATCCTCGC 1200
Db 181 GGTGGACCCGCTCAAGGAGCGCGGTGGTGAACCGCAACCGTCTCTCCATCCTCGC 240
Qy 1201 CGTAGACTACCCGTGACAAAGTCTCTGCTAGTCTCCGACGCGCGGTGATGCT 1260
Db 241 CGTAGACTACCCGTGACAAAGTCTCTGCTAGTCTCCGACGCGCGGTGATGCT 300
Qy 1261 GACGTTTCGAGTCGCTGTCGAGAGCGCGCGGTGTCGCGCAAGTGGTGGCCCTTCGCA 1320
Db 301 GACGTTTCGAGTCGCTGTCGAGAGCGCGCGGTGTCGCGCAAGTGGTGGCCCTTCGCA 360
Qy 1321 GAAGTTCGGCATCGAGCCCGCGCGCGCGGTGTCGCTTCGCTCAAGGTCGACTACCT 1380
Db 361 GAAGTTCGGCATCGAGCCCGCGCGCGCGGTGTCGCTTCGCTCAAGGTCGACTACCT 420
Qy 1381 CAAGGACAAAGTGCAGCCACCTTCTGTCGAGG 1412
Db 421 CAAGGACAAAGTGCAGCCACCTTCTGTCGAGG 452

RESULT 16
CA183623
LOCUS
DEFINITION
CA183623 608 bp mRNA linear EST 24-SEP-2003
SCQGST3154G05.g ST3 Saccharum officinarum cDNA clone SCQGST3154G05
5', mRNA sequence.

ACCESSION
CA183623
VERSION
CA183623.1 GI:35120577
KEYWORDS
EST.
SOURCE
Saccharum officinarum
ORGANISM
Saccharum officinarum

REFERENCE
1 (bases 1 to 608)
Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenhariaia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089

Email: parruda@unicamp.br
Clone distribution: Clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 154 row: G column: 05
Seq primer: T7 Promoter Primer.
Location/Qualifiers

FEATURES

source
1. .608
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCQST3154G05"
/lab_host="DH10B"
/clone_lib="ST3"
/notes="Organ: Fourth apical stalk internodes of adult
plants; Vector: pSport1; Site 1: SalI; Site 2: NotI; An
unidirectional cDNA library generated from [Fourth apical
stalk internodes of adult plants]. cDNA was prepared from
polyA+ mRNA using SuperScript Plasmid System Kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Query Match 9.3%; Score 320; DB 6; Length 608;
Best Local Similarity 100.0%; Pred. No. 5.8e-139; Indels 0; Gaps 0;
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 TCGCACAAACCGAAGCGAGCTGGTGTGATCCGGGGCCACGAGGACCCCAAGCGCTGCGG 181
Db 152 TCGCACAAACCGAAGCGAGCTGGTGTGATCCGGGGCCACGAGGACCCCAAGCGCTGCGG 211
QY 182 GCCTGAGCGGGGAGGTGTCGAGATATCGGCGACGAGGTCCGGCTCAGGTGGACGGC 241
Db 212 GCCTGAGCGGGGAGGTGTCGAGATATCGGCGACGAGGTCCGGCTCAGGTGGACGGC 271
QY 242 GACCTCTTCCTGCTGCAACGAGTGGCTTCCCGTGTGCGGGCCCTGCTACGAGTAC 301
Db 272 GACCTCTTCCTGCTGCAACGAGTGGCTTCCCGTGTGCGGGCCCTGCTACGAGTAC 331
QY 302 GAGCGCGGAGGACGCAAGACTGCGCCCATGCAAGACGCGCTACAGCGCTCAAG 361
Db 332 GAGCGCGGAGGACGCAAGACTGCGCCCATGCAAGACGCGCTACAGCGCTCAAG 391
QY 362 GGGAGCCGAGGCTTCCCGGGACGATGACGAGGAGGACATCGACGACCTGGAGCGAG 421
Db 392 GGGAGCCGAGGCTTCCCGGGACGATGACGAGGAGGACATCGACGACCTGGAGCGAG 451
QY 422 TTCAACATCGACGACGAGAA 441
Db 452 TTCAACATCGACGACGAGAA 471

RESULT 17
CA173049 624 bp mRNA linear EST 24-SEP-2003
LOCUS SCUTSB1033C12.g SBI Saccharum officinarum cDNA clone SCUTSB1033C12
5', mRNA sequence.
CA173049
CA173049.1 GI:35099270
EST.

ORGANISM Saccharum officinarum
Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE 1 (bases 1 to 624)
Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
The libraries that made SUCEST

JOURNAL COMMENT

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 033 row: C column: 12
Seq primer: T7 Promoter Primer.
Location/Qualifiers

FEATURES

source
1. .624
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCUTSB1033C12"
/lab_host="DH10B"
/clone_lib="SBI"
/notes="Organ: Stalk Bark from adult plants; Vector:
pSport1; Site 1: SalI; Site 2: NotI; An unidirectional
cDNA library generated from [Stalk Bark from adult
plants]. cDNA was prepared from polyA+ mRNA using
SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Query Match 9.3%; Score 320; DB 6; Length 624;
Best Local Similarity 100.0%; Pred. No. 5.8e-139; Indels 0; Gaps 0;
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 TCGCACAAACCGAAGCGAGCTGGTGTGATCCGGGGCCACGAGGACCCCAAGCGCTGCGG 181
Db 125 TCGCACAAACCGAAGCGAGCTGGTGTGATCCGGGGCCACGAGGACCCCAAGCGCTGCGG 184
QY 182 GCCTGAGCGGGGAGGTGTCGAGATATCGGCGACGAGGTCCGGCTCAGGTGGACGGC 241
Db 185 GCCTGAGCGGGGAGGTGTCGAGATATCGGCGACGAGGTCCGGCTCAGGTGGACGGC 244
QY 242 GACCTCTTCCTGCTGCAACGAGTGGCTTCCCGTGTGCGGGCCCTGCTACGAGTAC 301
Db 245 GACCTCTTCCTGCTGCAACGAGTGGCTTCCCGTGTGCGGGCCCTGCTACGAGTAC 304
QY 302 GAGCGCGGAGGACGCAAGACTGCCCCCATGCAAGACGCGCTACAGCGCTCAAG 361
Db 305 GAGCGCGGAGGACGCAAGACTGCCCCCATGCAAGACGCGCTACAGCGCTCAAG 364
QY 362 GGGAGCCGAGGCTTCCCGGGACGATGACGAGGAGGACATCGACGACCTGGAGCGAG 421
Db 365 GGGAGCCGAGGCTTCCCGGGACGATGACGAGGAGGACATCGACGACCTGGAGCGAG 424
QY 422 TTCAACATCGACGACGAGAA 441
Db 425 TTCAACATCGACGACGAGAA 444

RESULT 18 CG436514 LOCUS

DEFINITION CG436514 808 bp DNA linear GSS 17-SEP-2003
CG5EK86TV ZM.0.7_1.5_KB Zea mays genomic clone ZMMEMa0846P03,
genomic survey sequence.
CG436514
CG436514.1 GI:34813053
GSS.
KEYWORDS Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

TITLE
JOURNAL
COMMENT

Unpublished (2002)

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TP

Class: sheared ends.

Location/Qualifiers

1. .808

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBma0846p03"

/clone_lib="ZM 0.7 1.5 KB"

/note="Vector: pBSK-; Site 1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"

ORIGIN

Query Match

Best Local Similarity

Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Score 318; DB 9; Length 808;

Pred. No. 5e-138;

Qy

46 GGGTCGAGGCGGAGAGTACCTAGCAGCGCGCTCCGCGCGATGGAGGCCAGCGC 105

Db

1 GGGTCGAGGCGGAGAGTACCTAGCAGCGCGCTCCGCGCGATGGAGGCCAGCGC 60

Qy

106 CGGCTGTGTGCGCGCTCGCACAACCGAGCTGTGTGTATCGGGCGCAGGGA 165

Db

61 CGGCTGTGTGCGCGCTCGCACAACCGAGCTGTGTGTATCGGGCGCAGGGA 120

Qy

166 CCCCAGCGCTGCGGGCGCTGAGCGGCGAGTGTGCGAGATATGCGGCGCAGGTCGG 225

Db

121 CCCCAGCGCTGCGGGCGCTGAGCGGCGAGTGTGCGAGATATGCGGCGCAGGTCGG 180

Qy

226 GCTCACGCTGACGCGGACCTCTTGTGCTGCTGCAACAGAGTCGCGCTTCCCGTGGCG 285

Db

181 GCTCACGCTGACGCGGACCTCTTGTGCTGCTGCAACAGAGTCGCGCTTCCCGTGGCG 240

Qy

286 GCGCTGTACAGTACGAGCGCGGAGGCGCAGCACTGCCCCCAGTCAAGAGCGC 345

Db

241 GCGCTGTACAGTACGAGCGCGGAGGCGCAGCACTGCCCCCAGTCAAGAGCGC 300

Qy

346 CTACAAGCGCTCAAGGG 363

Db

301 CTACAAGCGCTCAAGGG 318

RESULT 19

CG454501

LOCUS

DEFINITION

CG454501

genomic survey sequence.

CG454501

VERSION

CG454501.1

GI:34839501

KEYWORDS

GSS.

SOURCE

ORGANISM

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 748)

REFERENCE

AUTHORS

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,

Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and

10-627-132-29.olg25.rst

10-627-132-29.olg25.rst

10-627-132-29.olg25.rst

10-627-132-29.olg25.rst

10-627-132-29.olg25.rst

10-627-132-29.olg25.rst

10-627-132-29.olg25.rst

10-627-132-29.olg25.rst

10-627-132-29.olg25.rst

10-627-132-29.olg25.rst

TITLE
JOURNAL
COMMENT

Unpublished (2003)

Other GSSs: PUIKA96TD PUIKA96TDB

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1. .748

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBma0601023"

/clone_lib="ZM 0.6 1.0 KB"

/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high

Cot selected genomic DNA library"

ORIGIN

Query Match

Best Local Similarity

Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Score 305; DB 9; Length 748;

Pred. No. 6.6e-132;

Qy

2307 AGCTGGGGTGGATCTACGGGTTCGATCAGGAGGACATCTTACGGGGTTCAAGATGCACT 2366

Db

149 AGCTGGGGTGGATCTACGGGTTCGATCAGGAGGACATCTTACGGGGTTCAAGATGCACT 208

Qy

2367 GCGCGGGTGGCGTCCGTGTACTGTGATGCCAGCGGCGCGCTTCAAGGGGTTCGGCGC 2426

Db

209 GCGCGGGTGGCGTCCGTGTACTGTGATGCCAGCGGCGCGCTTCAAGGGGTTCGGCGC 268

Qy

2427 CGATCAATCTATCGGACCGTCTCAACAGGTCTCCGGTGGCGGCTCCGTCGAGCA 2486

Db

269 CGATCAATCTATCGGACCGTCTCAACAGGTCTCCGGTGGCGGCTCCGTCGAGCA 328

Qy

2487 TCTTCTTTCAGCGGACAGCCCCCTGTGTACGGCTACAAGACGGCAACCTCAAGTGGC 2546

Db

329 TCTTCTTTCAGCGGACAGCCCCCTGTGTACGGCTACAAGACGGCAACCTCAAGTGGC 388

Qy

2547 TGGAGCGCTTTCGCTTACATCAACACCACTTACCCCTTACCTCGCTCCCGTGTCTG 2606

Db

389 TGGAGCGCTTTCGCTTACATCAACACCACTTACCCCTTACCTCGCTCCCGTGTCTG 448

Qy

2607 CTTACTGCACCTTCCCGCGCTGTGCTCTCTCACCAGGCAAGTTTCATATGCCGTCG 2662

Db

449 CTTACTGCACCTTCCCGCGCTGTGCTCTCTCACCAGGCAAGTTTCATATGCCGTCG 504

RESULT 20

CC705263

LOCUS

DEFINITION

OGWEX08TH ZM 0.7 1.5 KB Zea mays genomic clone ZMMBma0562B16,

genomic survey sequence.

CC705263

VERSION

CC705263.1

GI:32110039

KEYWORDS

GSS.

SOURCE

ORGANISM

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 640)

REFERENCE

AUTHORS

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,

Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,

Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

TITLE

JOURNAL

COMMENT

Other_GSSs: OGWEX08TV

10-627-132-29.olg25.rst

10-627-132-29.olg25.rst

10-627-132-29.olg25.rst

10-627-132-29.olg25.rst

10-627-132-29.olg25.rst

10-627-132-29.olg25.rst

10-627-132-29.olg25.rst

10-627-132-29.olg25.rst

Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES

source
1. .640
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0562B16"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCSk-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 7.8%; Score 269; DB 9; Length 640;
Best Local Similarity 100.0%; Pred. No. 5.9e-115;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 817 GAACGACGAGCGAGGAGCGGCTGTCGAGGAAGGTTCGATCGCTCGAGCAAGGTGAA 876
Db 298 GAACGACGAGCGAGGAGCGGCTGTCGAGGAAGGTTCGATCGCTCGAGCAAGGTGAA 357
QY 877 CCGGTACCGGATGATCGTGGTGGCTGCTCGTGTGCTCGCTTCTTCCTCCGATACCG 936
Db 358 CCGGTACCGGATGATCGTGGTGGCTGCTCGTGTGCTCGCTTCTTCCTCCGATACCG 417
QY 937 TATCCTGCACCCCGTCCCGAGCGCATCGGCGCTGTGCTGTGCTTCCATCATCTGCGAGAT 996
Db 418 TATCCTGCACCCCGTCCCGAGCGCATCGGCGCTGTGCTGTGCTTCCATCATCTGCGAGAT 477
QY 997 CTGGTTCGCCATCTCTCGATCTCTGATCTCCAGTTCCTCCCAAGTGGTTCCTCCATCGACCGCA 1056
Db 478 CTGGTTCGCCATCTCTCGATCTCTGATCTCCAGTTCCTCCCAAGTGGTTCCTCCATCGACCGCA 537
QY 1057 GAGTACTCGACCGCTCTCCCTCAGGT 1085
Db 538 GAGTACTCGACCGCTCTCCCTCAGGT 566

RESULT 21

CC446976 833 bp DNA linear GSS 20-MAY-2003
LOCUS
DEFINITION
PURLI43TD ZM 0.6 1.0 KB Zea mays genomic clone ZMMBma477H13,
genomic survey sequence.
ACCESSION
VERSION CC446976.1 GI:30954669
KEYWORDS
SOURCE
GSS.
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS
1 (bases 1 to 833)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Reinick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Bennetzen, J.
TITLE
Maize Genomics Consortium
JOURNAL
COMMENT
Unpublished (2003)
Other GSSs: PURLI43TB
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.

FEATURES

source
1. .833
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma477H13"
/clone_lib="ZM 0.6 1.0 KB"
/notes="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN

Query Match 7.8%; Score 269; DB 8; Length 833;
Best Local Similarity 100.0%; Pred. No. 5.8e-115;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 817 GAACGACGAGCGAGGAGCGGCTGTCGAGGAAGGTTCGATCGCTCGAGCAAGGTGAA 876
Db 130 GAACGACGAGCGAGGAGCGGCTGTCGAGGAAGGTTCGATCGCTCGAGCAAGGTGAA 189
QY 877 CCGGTACCGGATGATCGTGGTGGCTGCTCGTGTGCTCGCTTCTTCCTCCGATACCG 936
Db 190 CCGGTACCGGATGATCGTGGTGGCTGCTCGTGTGCTCGCTTCTTCCTCCGATACCG 249
QY 937 TATCCTGCACCCCGTCCCGAGCGCATCGGCGCTGTGCTGTGCTTCCATCATCTGCGAGAT 996
Db 250 TATCCTGCACCCCGTCCCGAGCGCATCGGCGCTGTGCTGTGCTTCCATCATCTGCGAGAT 309
QY 997 CTGGTTCGCCATCTCTCGATCTCTGATCTCCAGTTCCTCCCAAGTGGTTCCTCCATCGACCGCA 1056
Db 310 CTGGTTCGCCATCTCTCGATCTCTGATCTCCAGTTCCTCCCAAGTGGTTCCTCCATCGACCGCA 369
QY 1057 GAGTACTCGACCGCTCTCCCTCAGGT 1085
Db 370 GAGTACTCGACCGCTCTCCCTCAGGT 398

RESULT 22

BZ733711/c
LOCUS
DEFINITION
OGEDM54TW ZM 0.7 1.5 KB Zea mays genomic clone ZMMBma0251J11,
genomic survey sequence.
ACCESSION
VERSION BZ733711.1 GI:28709849
KEYWORDS
SOURCE
GSS.
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS
1 (bases 1 to 841)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Reinick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
TITLE
Consortium for Maize Genomics
JOURNAL
COMMENT
Unpublished (2002)
Other GSSs: OGEDM54TC
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
source
1. .841
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0251J11"
/clone_lib="ZM 0.7 1.5 KB"

Qy	1488	TGCCGGCAGAGGGGTGATCATGAAGAACCGCAGCCGTGGCCGGGAACAACACCCGCG	1547
Db <td>376</td> <td>TGCCGGCAGAGGGGTGATCATGAAGAACCGCAGCCGTGGCCGGGAACAACACCCGCG</td> <td>317</td>	376	TGCCGGCAGAGGGGTGATCATGAAGAACCGCAGCCGTGGCCGGGAACAACACCCGCG	317

```
QY 1548 ACCACCCCGCATGATCCAGGTGTTCTTGGGCCACAGCGCGGCCACGACACCGAGGGCA 1607
Db 316 ACCACCCCGCATGATCCAGGTGTTCTTGGGCCACAGCGCGGCCACGACACCGAGGGCA 257
QY 1608 ACCAGCTGCCCGCCTCGTCTAGCTCTCCGTCGAGAGCGCCCGGATTCACGACACCA 1667
Db 256 ACAGAGCTGCCCGCCTCGTCTAGCTCTCCGTCGAGAGCGCCCGGATTCACGACACCA 197
QY 1668 AGAAGCGCGCGCATGAACGCTCTG 1693
Db 196 AGAAGCGCGCGCATGAACGCTCTG 171

RESULT 25
CC656942/c
LOCUS
DEFINITION
OGLAD377H ZM 0.7_1.5_KB Zea mays genomic clone ZMBMa0307H02,
genomic survey sequence.
ACCESSION
CC656942
VERSION
CC656942.1 GI:32060234
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 818)
AUTHORS
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OGLAD37TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
Location/Qualifiers
source
1..818
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBMa0307H02"
/notes="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 7.7%; Score 266; DB 9; Length 818;
Best Local Similarity 100.0%; Pred. No. 1.5e-113;
Matches 266; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1428 AGAGAGGATGAGGAGTTCAAGGTCGGATCAACGCGCTGTGGCCAGGCCATGAAGG 1487
Db 542 AGAGAGGATGAGGAGTTCAAGGTCGGATCAACGCGCTGTGGCCAGGCCATGAAGG 483
QY 1488 TGCCGGCAGAGGGGTGGATCATGAAGGACGGCACGCGCTGCCCGGGGAACAACACCCGCG 1547
Db 482 TGCCGGCAGAGGGGTGGATCATGAAGGACGGCACGCGCTGCCCGGGGAACAACACCCGCG 423
QY 1548 ACCACCCCGCATGATCCAGGTGTTCTTGGGCCACAGCGCGGCCACGACACCGAGGGCA 1607
Db 422 ACCACCCCGCATGATCCAGGTGTTCTTGGGCCACAGCGCGGCCACGACACCGAGGGCA 363
QY 1608 ACCAGCTGCCCGCCTCGTCTAGCTCTCCGTCGAGAGCGCCCGGATTCACGACACCA 1667
Db 362 ACCAGCTGCCCGCCTCGTCTAGCTCTCCGTCGAGAGCGCCCGGATTCACGACACCA 303
```

```
QY 1668 AGAAGCGCGCGCATGAACGCTCTG 1693
Db 302 AGAAGCGCGCGCATGAACGCTCTG 277

RESULT 26
CG343883
LOCUS
DEFINITION
OG0AF68TV ZM 0.7_1.5_KB Zea mays genomic clone ZMBMa0672K16,
genomic survey sequence.
ACCESSION
CG343883
VERSION
CG343883.1 GI:34261149
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 868)
AUTHORS
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OG0AF68TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
Location/Qualifiers
source
1..868
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBMa0672K16"
/clone_lib="ZM_0.7_1.5_KB"
/notes="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 7.7%; Score 266; DB 9; Length 868;
Best Local Similarity 100.0%; Pred. No. 1.5e-113;
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1428 AGAGAGGATGAGGAGTTCAAGGTCGGATCAACGCGCTGTGGCCAGGCCATGAAGG 1487
Db 116 AGAGAGGATGAGGAGTTCAAGGTCGGATCAACGCGCTGTGGCCAGGCCATGAAGG 175
QY 1488 TGCCGGCAGAGGGGTGGATCATGAAGGACGGCACGCGCTGCCCGGGGAACAACACCCGCG 1547
Db 176 TGCCGGCAGAGGGGTGGATCATGAAGGACGGCACGCGCTGCCCGGGGAACAACACCCGCG 235
QY 1548 ACCACCCCGCATGATCCAGGTGTTCTTGGGCCACAGCGCGGCCACGACACCGAGGGCA 1607
Db 236 ACCACCCCGCATGATCCAGGTGTTCTTGGGCCACAGCGCGGCCACGACACCGAGGGCA 295
QY 1608 ACCAGCTGCCCGCCTCGTCTAGCTCTCCGTCGAGAGCGCCCGGATTCACGACACCA 1667
Db 296 ACCAGCTGCCCGCCTCGTCTAGCTCTCCGTCGAGAGCGCCCGGATTCACGACACCA 355
QY 1668 AGAAGCGCGCGCATGAACGCTCTG 1693
Db 356 AGAAGCGCGCGCATGAACGCTCTG 381

RESULT 27
CG145361/c
LOCUS
DEFINITION
OG145361 ZM 0.7_1.5_KB Zea mays genomic clone ZMBMa0672K16,
genomic survey sequence.
ACCESSION
CG145361
VERSION
CG145361.1 GI:34261149
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 868)
AUTHORS
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OG0AF68TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
Location/Qualifiers
source
1..868
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBMa0672K16"
/clone_lib="ZM_0.7_1.5_KB"
/notes="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 7.7%; Score 266; DB 9; Length 868;
Best Local Similarity 100.0%; Pred. No. 1.5e-113;
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1428 AGAGAGGATGAGGAGTTCAAGGTCGGATCAACGCGCTGTGGCCAGGCCATGAAGG 1487
Db 116 AGAGAGGATGAGGAGTTCAAGGTCGGATCAACGCGCTGTGGCCAGGCCATGAAGG 175
QY 1488 TGCCGGCAGAGGGGTGGATCATGAAGGACGGCACGCGCTGCCCGGGGAACAACACCCGCG 1547
Db 176 TGCCGGCAGAGGGGTGGATCATGAAGGACGGCACGCGCTGCCCGGGGAACAACACCCGCG 235
QY 1548 ACCACCCCGCATGATCCAGGTGTTCTTGGGCCACAGCGCGGCCACGACACCGAGGGCA 1607
Db 236 ACCACCCCGCATGATCCAGGTGTTCTTGGGCCACAGCGCGGCCACGACACCGAGGGCA 295
QY 1608 ACCAGCTGCCCGCCTCGTCTAGCTCTCCGTCGAGAGCGCCCGGATTCACGACACCA 1667
Db 296 ACCAGCTGCCCGCCTCGTCTAGCTCTCCGTCGAGAGCGCCCGGATTCACGACACCA 355
QY 1668 AGAAGCGCGCGCATGAACGCTCTG 1693
Db 356 AGAAGCGCGCGCATGAACGCTCTG 381

RESULT 27
CG145361/c
LOCUS
DEFINITION
OG145361 ZM 0.7_1.5_KB Zea mays genomic clone ZMBMa0672K16,
genomic survey sequence.
ACCESSION
CG145361
VERSION
CG145361.1 GI:34261149
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 868)
AUTHORS
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OG0AF68TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
Location/Qualifiers
source
1..868
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBMa0672K16"
/clone_lib="ZM_0.7_1.5_KB"
/notes="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 7.7%; Score 266; DB 9; Length 868;
Best Local Similarity 100.0%; Pred. No. 1.5e-113;
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1428 AGAGAGGATGAGGAGTTCAAGGTCGGATCAACGCGCTGTGGCCAGGCCATGAAGG 1487
Db 116 AGAGAGGATGAGGAGTTCAAGGTCGGATCAACGCGCTGTGGCCAGGCCATGAAGG 175
QY 1488 TGCCGGCAGAGGGGTGGATCATGAAGGACGGCACGCGCTGCCCGGGGAACAACACCCGCG 1547
Db 176 TGCCGGCAGAGGGGTGGATCATGAAGGACGGCACGCGCTGCCCGGGGAACAACACCCGCG 235
QY 1548 ACCACCCCGCATGATCCAGGTGTTCTTGGGCCACAGCGCGGCCACGACACCGAGGGCA 1607
Db 236 ACCACCCCGCATGATCCAGGTGTTCTTGGGCCACAGCGCGGCCACGACACCGAGGGCA 295
QY 1608 ACCAGCTGCCCGCCTCGTCTAGCTCTCCGTCGAGAGCGCCCGGATTCACGACACCA 1667
Db 296 ACCAGCTGCCCGCCTCGTCTAGCTCTCCGTCGAGAGCGCCCGGATTCACGACACCA 355
QY 1668 AGAAGCGCGCGCATGAACGCTCTG 1693
Db 356 AGAAGCGCGCGCATGAACGCTCTG 381
```

DEFINITION PUJEW43TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa0652H13,
genomic survey sequence.
ACCESSION CG145361
VERSION CG145361.1 GI:34036144
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 795)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other GSSs: PUJEW43TB
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: Tg
Class: Sheared ends.
FEATURES
source
1. .795
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa0652H13"
/clone_lib="ZM 0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
CoT selected genomic DNA library"
ORIGIN
Query Match 7.6%; Score 261; DB 9; Length 795;
Best Local Similarity 100.0%; Pred. No. 3.4e-111;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3146 TCCATTCTGCTGGCTCCATCTTCTCCTGCTCTGGGTGAGATCGACCTTCATCGTC 3205
DB 795 TCCATTCTGCTGGCTCCATCTTCTCCTGCTCTGGGTGAGATCGACCTTCATCGTC 736
QY 3206 AGGACCAAGGCGCGGAGCTGAGGAGTGGCATCAATTGCTGAGCTGTTTATTAAAGT 3265
DB 735 AGGACCAAGGCGCGGAGCTGAGGAGTGGCATCAATTGCTGAGCTGTTTATTAAAGT 676
QY 3266 TCAAAATTCGAGCTTGTGATAGGGAGAAAAAACAATTAGAAATTTTGTAAAGTTG 3325
DB 675 TCAAAATTCGAGCTTGTGATAGGGAGAAAAAACAATTAGAAATTTTGTAAAGTTG 616
QY 3326 TTGTGCTGTAATGTTATGTTACCAATGTCGAGGAGGAATTTGAACAAAGGACAAG 3385
DB 615 TTGTGCTGTAATGTTATGTTACCAATGTCGAGGAGGAATTTGAACAAAGGACAAG 556
QY 3386 GTTTGATTGTTAAATGGCAA 3406
DB 555 GTTTGATTGTTAAATGGCAA 535
RESULT 28
BE050992
LOCUS BE050992
DEFINITION za71d11.g51 Maize Glume cDNAs Library Zea mays cDNA clone za71d11
5', mRNA sequence.
ACCESSION BE050992
VERSION BE050992.1 GI:8368047
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 437)
AUTHORS O'Shaughnessy,A.L., Habermann,K., de la Bastide,M., Huang,E.N.,
Nascimento,L.U., Schutz,K., Matero,A., Swaby,I., See,L.-H.,
Preston,R.R., Rodriguez,M.A., Shah,R.S., Shekher,M., Spiegel,L.A.,
Vil,M.D., Dedhia,N.N. and McCombie,W.R.
Expressed sequence tags from Zea mays (maize)
Unpublished (2000)
COMMENT Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: za71 row: d column: 11
Seq primer: -40M13RevUniv
High quality sequence stop: 437.
FEATURES
source
1. .437
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
/clone="za71d11"
/clone_lib="Maize Glume cDNAs Library"
/note="Vector: Lambda Zap II (Stratagene); Site 1: XhoI;
Site 2: EcoRI; Resistance: Ampicillin; Autoexcision:
pBluescript SK (+/-); Titer: 7 x 10e-9 pfu/mL (as of
9/28/94); Library source: John Doebley, B73"
ORIGIN
Query Match 7.1%; Score 246; DB 2; Length 437;
Best Local Similarity 100.0%; Pred. No. 4.1e-104;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1778 CGGAGGCCATGCTTCTCTCATGACCCCTCAGTCGGCGGAGGTCTGCTACGTTTCAG 1837
DB 192 CGGAGGCCATGCTTCTCTCATGACCCCTCAGTCGGCGGAGGTCTGCTACGTTTCAG 251
QY 1838 TTCCCGCAGAGGTTTCGACCGCATCGAGTCGACGACGATACGTAACAGGACACCGTC 1897
DB 252 TTCCCGCAGAGGTTTCGACCGCATCGAGTCGACGACGATACGTAACAGGACACCGTC 311
QY 1898 TTCTTCGACATCAACATGAAGGGCTCGACGGCATCCAAGCCCGGTGTAGTCGGGACA 1957
DB 312 TTCTTCGACATCAACATGAAGGGCTCGACGGCATCCAAGCCCGGTGTAGTCGGGACA 371
QY 1958 GGGTGGCTGTTCGGCGCCAGGGCTCTACGGCTAGAACCTCCCAAGGACCCCAAGG 2017
DB 372 GGGTGGCTGTTCGGCGCCAGGGCTCTACGGCTAGAACCTCCCAAGGACCCCAAGG 431
QY 2018 CCCAAG 2023
DB 432 CCCAAG 437
RESULT 29
BZ733703
LOCUS BZ733703
DEFINITION OGEDM54TC ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0251J11,
genomic survey sequence.
ACCESSION BZ733703
VERSION BZ733703.1 GI:28709833
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 821)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 12, 2005, 10:16:14 ; Search time 5122 Seconds
(without alignments)
7817.964 Million cell updates/sec

Title: US-10-627-132-30

Perfect score: 5648

Sequence: 1 MEASAGLVAGSHNRNELVLI.....IDPFIVRTKGPDPVRCQGINC 1052

Scoring table:

BLOSUM62
-Q=/cg2 1/USPTO spool-US10627132/runat 04032005 113218 21160/app query.fasta_1.1223
-DB=EST -QFMT=fastap -SUPFIX=est -MINMATCH=0.1 -FOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=spt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-USER=US10627132 @CGN 1 1 3043 @runat 04032005 113218 21160 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cg2 1/USPTO spool-US10627132/runat 04032005 113218 21160/app query.fasta_1.1223
-DB=EST -QFMT=fastap -SUPFIX=est -MINMATCH=0.1 -FOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=spt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-USER=US10627132 @CGN 1 1 3043 @runat 04032005 113218 21160 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb est1:*
2: gb est2:*
3: gb htc:*
4: gb est3:*
5: gb est4:*
6: gb est5:*
7: gb est6:*
8: gb gsa1:*
9: gb gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5347.5	94.7	3168	9	CL976367
2	4089	72.4	3783	3	AY104730
3	4061.5	71.9	3897	3	AY103655
4	3938	69.7	3788	3	AY103701
5	3912.5	69.3	3763	3	AY108113
6	3907	69.2	3192	9	CL972423
7	3821.5	67.7	3898	3	AY110415
8	3719	65.8	3911	3	CNS0A1ML
9	3633.5	64.3	3728	3	AY112236

10	3597.5	63.7	3696	3	AY110079
11	3350.5	59.3	2872	3	AY104236
12	2525	44.7	1874	6	CD726831
13	2311	40.9	3474	9	CL978864
14	2134.5	37.8	2595	9	CL964957
15	1982	35.1	3048	9	CL974686
16	1607.5	28.5	2640	9	CL973331
17	1599.5	28.3	2670	9	CL973334
18	1549.5	27.4	2607	9	CL973335
19	1539.5	27.3	960	1	AB049602
20	1531	27.1	1797	9	CL979846
21	1434	25.4	898	7	CO089073
22	1432.5	25.4	1189	3	AY107656
23	1424	25.2	916	7	CK272603
24	1421	25.2	924	7	CF513822
25	1395	24.7	789	7	CF430024
26	1365	24.2	1152	3	CNS0A924
27	1362.5	24.1	925	7	CF213555
28	1344	23.8	953	6	CD439980
29	1302.5	23.1	848	7	CN129331
30	1275	22.6	821	9	CG383072
31	1272.5	22.5	853	7	CO086716
32	1272	22.5	875	7	CF514581
33	1268	22.5	869	7	CO089781
34	1264	22.4	779	6	CA916873
35	1262.5	22.4	837	7	CF512600
36	1262	22.3	834	7	CO519653
37	1258	22.3	792	6	CB671068
38	1249	22.1	772	6	CB646549
39	1246	22.1	790	6	CA161838
40	1246	22.1	851	6	CB892280
41	1240	22.0	799	6	CB894714
42	1239	21.9	819	7	CO111111
43	1237	21.9	710	6	CA261021
44	1236	21.9	868	9	CC660677
45					

ALIGNMENTS

RESULT 1	CL976367	3168 bp	DNA	linear	GSS 21-SEP-2004
LOCUS	CL976367	Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.			
DEFINITION	CL976367	GI:52407247			
ACCESSION	CL976367				
VERSION	CL976367.1				
KEYWORDS	GSS.				
SOURCE	Oryza sativa (indica cultivar-group)				
ORGANISM	Oryza sativa (indica cultivar-group)				
REFERENCE	Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Zhao, X., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G.K.S., Deng, X.W. and Wang, J.				
AUTHORS	An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis				
TITLE	Unpublished (2004)				
JOURNAL	Contact: Chen Chen				
COMMENT	Department of Bioinformatic				
	Beijing Institute of Genomics				
	Chinese Academy of Sciences, Beijing 101300, China				
	Tel: 86-10-80481559				
	Fax: 86-10-80488676				
	Email: chenchen@genomics.org.cn				
	Rice genomic sequence.				
	Class: exon-trapped.				
	Location/Qualifiers				
	1..3168				
	/organism="Oryza sativa (indica cultivar-group)"				
	/mol_type="genomic DNA"				

FEATURES
source

/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed Library"
/note="Oryza sativa exon trapped genomic sequences "

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3168
Score: 5347.50 Matches: 992
Percent Similarity: 96.78% Conservative: 30
Best Local Similarity: 93.94% Mismatches: 29
Query Match: 94.68% Indels: 5
DB: 9 Gaps: 4

US-10-627-132-30 (1-1052) x CL976367 (1-3168)

Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuLeu 20
Db 1 ATGGAGGCGAGCGCGGGCTGGTGGCGGGTCCGACAAACCGGAAACGAGCTGGTCTGATC 60
Qy 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGluLeuCys 40
Db 61 CGGGGGACAGAGGAGCCCAAGCCGCTGCGGGCGCTGAGCGGGCAGGTGTGGAGATATGC 120
Qy 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
Db 121 GCGCAGCAGGTCGCGCGCACCGTCGACGCGGACCTTCCTGTCGCTGCAACGAGTGGCGC 180
Qy 61 PheProValCysArgProCysTyrgluTyrgluArgArgGluGlyThrGlnAsnCysPro 80
Db 181 TTCGCGGTGTCCGCCCTGTCTAGAGTACGAGGCGCGCGAGGACCCCAAACTGGCCCC 240
Qy 81 GlnCysLysThrArgTyrgluArgLeuLysGlySerProArgValAlaGlyAspAsp 100
Db 241 CAGTGCAAGACCCGCTACAAGCGCTCAAGGGGAGCCCGAGGGTGCCCGGGGACGAGAC 300
Qy 101 GluGluAspLeuAspAspGluHisGluPheAsnLeuAspAspGluAsnGlnGln--- 119
Db 301 GAGGAGGACATTGACGACCTGGAGCAGCAGATTCAACATCGACGACGAGAACGACGAC 360
Qy 120 ---ArgGlnLeuGluGlyAsnMetGlnAsnSerGlnLeuThrGluAlaMetLeuHisGly 138
Db 361 CTGCAGCAGGATCAGGATGGCATGCAGAACAGCCACATCACCGAGGCGATCTGCACGGC 420
Qy 139 ArgMetSerTyrglyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnLeuPro 158
Db 421 AAGATGACTACGGAGGGGGCCCCGACGCGCGGCGGCAAGCAGCCCCG---CTCCCG 477
Qy 159 ProIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGly 178
Db 478 CCGATCATCCCGCGCTCGCTCGCTCGGTCGAGCGGGAGTTCCCGATATCGAACAGC 537
Qy 179 TyrGlyHisGlyGluValSerSerSerLeuHisLysArgIleHisProTyrglyProValSer 198
Db 538 CATGCCCATGGCGAGTTCTCTCTCTCCCTGCACAAAGCGCATCCACCCCTACCCCGGTCT 597
Qy 199 GluProGlySerAlaLysTrpAspGluLysLysGluValSerTrpLysGluArgMetAsp 218
Db 598 GAGCAGGGAGTGCAAGTGGGACGAGAAAGAGGTGAGCTGGAGGAGGATGGAC 657
Qy 219 AspTrpLysSerLysGlnGlyLeuGlyGlyAla---AspProGluAspMetAsp 237
Db 658 GACTGGAAATCCAACAGCGGCATCGTCGCGCGCGCGCCCCCGATCCCGACGACTACGAC 717
Qy 238 AlaAspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAla 257
Db 718 GCGGACGTCCTCACTGAACAGCAGGCGGAGCGCGCTGTCGAGGAGGTTGTCGATCGCG 777
Qy 258 SerSerLysValAsnProTyrglyArgMetValIleValValArgLeuValValAlaPhe 277
Db 778 TCGACCAAGGTGAACCCGTACCGGATGGTGATCATCTCCCTCGCTCTGCTGGTCTCGGCTTC 837
Qy 278 PheLeuArgTyrglyArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSer 297

Db 838 TTCCTCGGTACCGCATCTCTCCACCGGTGTCGCCAGCCATCCCGTGTGCTGCACCTCC 897
Qy 298 IleIleCysGluLeuTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPhe 317
Db 898 ATCATCTGGAGATCTGGTTCGCCGTGTCGTGATCTCTCACCAGTTCCCAAGTGTATC 957
Qy 318 ProIleAspArgGluThrTyrgluAspArgLeuSerLeuArgTyrgluArgGluGlyGlu 337
Db 958 CCGATCGACAGGGAGACCTACCTCGACCGCTCTCTCTCGCTACGAGCGCGAGGGGAG 1017
Qy 338 ProSerLeuLeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluPro 357
Db 1018 CCGTCTGCTATCGCGGTGGACCTGTCTCGTCAGCACGCTGATCTCGCTCAAGGAGCCG 1077
Qy 358 ProLeuValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrglyProValAspLys 377
Db 1078 CCGTGTGTACGGCCAAACCGTGTGTCATCTCTCGCTCGACTACCCCGTTCGACAG 1137
Qy 378 ValSerCysTyrglyValSerAspGlyAlaSerMetLeuThrPheGluSerLeuSerGlu 397
Db 1138 GTGTCTGTCTACGTCTCCGACGACGCGCTCCATCTCAGTTCGAGTCTGCTGCGAG 1197
Qy 398 ThrAlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArg 417
Db 1198 ACGGCGAGTTCGCCGCAAGTGGTGCCATTCTGCAAGAGTTTCAGCATCGAGCCCGC 1257
Qy 418 AlaProGluPheTyrglyPheSerLeuLysValAspTyrglyLeuLysAspLysValGlnProThr 437
Db 1258 GCGCGGAGTTCATTCTCCAGAGGTTCGATCTCAAGGACAGGTTCCATCCCAAC 1317
Qy 438 PheValGlnGluArgAlaMetLysArgGluTyrgluGluPheLysValArgIleAsn 457
Db 1318 TTCGTCCAGGCGCGCGCCATGAGAGAGAGTACGAGGAGTTCAAGGTGAGGATCAAC 1377
Qy 458 AlaLeuValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThr 477
Db 1378 GCGTGTGGCGAAGCGCAGAGGTTCGCCGCGAAGGTGGATCATGAAAGGACGGAGC 1437
Qy 478 ProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHis 497
Db 1438 CCATGGCGGGGAAACAAACCCCGCACCCCGGGCATGATCCAGGTGTTCTCGGCCAC 1497
Qy 498 SerGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrglyValSerArgGlu 517
Db 1498 AGCGCGGCCACGACCGAGGCGCAAGCTCCCGCGCTCGTCTACGTCTCCCGTCTGAG 1557
Qy 518 LysArgProGlyPheGlnHisLysLysLysAlaGlyAlaMetAsnAlaLeuLysArgVal 537
Db 1558 AAGCGCCCGGCTTCCAGCACCAAGAGCGCGCGCCATGAACGCGCTCATTTCTGTGTG 1617
Qy 538 SerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrglyIleAsn 557
Db 1618 TCGCGCTGTGTGACAAACCGCGCTTCATGCTCAACTTGGATTGCGATCACTATCATCAC 1677
Qy 558 AsnSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLys 577
Db 1678 AACAGCAAGGCCATCAGGAGGCGATGTGCTTCCTCATGGACCGCGAGTCCGACGGAAG 1737
Qy 578 ValCysTyrglyValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrgly 597
Db 1738 GTGTGCTACGTGCTGCTCCCGCAGAGGTTCGACGGCATCGACGCTCGACGCGGATACGCC 1797
Qy 598 AsnArgAsnThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyPro 617
Db 1798 AACCGCAACCCGTTCTTTCGACATCAACATGAAGGGGCTTGTGGGATCCAGGGCCCG 1857
Qy 618 ValTyrglyValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrglyTyrglyAsnProPro 637
Db 1858 GTGTACGTCCGGCAGGCGTGGTTCAGGCGCGAGCGCTGTACCGGATACCAACCCCTCCC 1917
Qy 638 LysGlyProLysArgProLysMetValThrCysAspCysCysProCysPheGlyArgLys 657
Db 1918 AAGGACCAAGAGGCCCAAGATGTTGACCTGCGACTGCTGCTGCTTGTCTCGGGAGGAG 1977


```
QY 658 LysArgLysHisAlaLysAspGlyLeuProGluGly---ThrAlaAspMetGlyValAsp 676
Db 1978 AAGCGGAAGACCGGAAGACGGCTCCCGGAGGGCGTCCGCGGAGCGGCGGATGGAC 2037
QY 677 SerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAla 696
Db 2038 AGCGACAGGAGATGCTCATGTCGAGATGAATCTCGAGAGCGGTTCGGGAGTCGGCG 2097
QY 697 AlaPheValThrSerThrLeuMetGluGluGlyValProProSerSerSerProAla 716
Db 2098 GCGTTCGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGAC 2157
QY 717 AlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrp 736
Db 2158 GCGCTCCTCAAGAGGCCATCATGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2217
QY 737 GlyLeuGluLeuGlyTyrPheLysSerIleThrGluAspIleLeuThrGlyPheLys 756
Db 2218 GGTCTCGAGCTGGGTGGATCTACGGTTCGATCAGGAGGATCTTACCGGGGTTCAG 2277
QY 757 MetHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGly 776
Db 2278 ATGCACTGCCGGGTGGAGATCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2337
QY 777 SerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySer 796
Db 2338 TCGGCGCGGATCAACCTATCTGACGCTCTCAACGAGGTCTCGGTGGGTCTCGGCTCC 2397
QY 797 ValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeu 816
Db 2398 GTGAGATCTCTTCAGCGCGGACAGCGCGCTCTCTACGGGTCTAAGACGCGAACCTT 2457
QY 817 LysTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuPro 836
Db 2458 AAGTGGCTCGAGCGTCTCTTACATCAACACCATCTACCTCTTCACTCTCTCTCTCC 2517
QY 837 LeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetPro 856
Db 2518 CTCTCTCGCTACTGCACTCTCTCGCGGCTCTGCTCTCTCTCTCTCTCTCTCTCTCT 2577
QY 857 SerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThr 876
Db 2578 CGATTACAGCTTTGCGAGTTGTCTTCATCGCGCTCTTCATCTCTCTCTCTCTCTCT 2637
QY 877 GlyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGln 896
Db 2638 GGCATCTCGAGATGAGTGGAGCGGGGTGAGCATCGAGGAGTGTGGAGGACGAGCAG 2697
QY 897 PheTrpValIleGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLys 916
Db 2698 TTCTGGGTCTATCGCGGGGTGTTCGGCGCACCTGTTCGGGTGTTCGAGGGCTGTCTCA 2757
QY 917 ValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAsp 936
Db 2758 GTGCTGGCGGGATCGACACCACTTCACCGTCAGTCCAGGCGCCACCGGAGACGAGAC 2817
QY 937 AspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThr 956
Db 2818 GACGAGTTCCGCGAGCTATACGCTTCAAGTGGACCACTCTCTCTCTCTCTCTCTCTCT 2877
QY 957 LeuLeuIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGly 976
Db 2878 CTGCTCATCTCAACATCATCGCGCTCTGCGCGCGCTCTCGAGCGCCATCAACACGCG 2937
QY 977 TyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHis 996
Db 2938 TCCGAGGCGTGGGGCGGCTCTTTCGGGAAGTCTTCTTCTGCGCTCTCTGCTCTCTCT 2997
QY 997 LeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgTrpProThrValValVal 1016
Db 2998 CTCTACCTCTCTCAAGGGGCTCATGGGAGGAGACCGGACCGCCACCATCTGTTGTC 3057
```

```
QY 1017 IleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProphe 1036
Db 3058 ATCTGGTCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3117
QY 1037 IleValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3118 ACCATCAAGCCAGGGGCGCTGACGTGAGCGAGTGGGATCACTGC 3165
RESULT 2
LOCUS AY104730 3783 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays PCO100501 mRNA sequence.
ACCESSION AY104730
VERSION AY104730.1 GI:21207808
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 3783)
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitesitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 3783)
Coe, E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
source
1. 3783
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635765"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BAGs in conjunction with the Maize
Mapping Project"
```

```
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3783
Score: 4089.00 Matches: 760
Percent Similarity: 79.55% Conservative: 127
Best Local Similarity: 68.16% Mismatches: 138
Query Match: 72.40% Indels: 90
DB: 3 Gaps: 16
US-10-627-132-30 (1-1052) x AY104730 (1-3783)
```

```
QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
Db 186 ATGAGGCGGAGCGCGGGCTGTGCGCGCTGCGCACACCGGAGAGCTGTGCTGTCATC 245
QY 21 Arg-----GlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGlu 38
Db 246 CGCCGCGATGCGAGCGGCGGCGGAGCCCATGGCCAGCGGCGGCGGCGGCGGCGG 305
QY 39 IleCysGlyAspGluValGlyLeuThrValAspGlyLeuValAlaCysAsnGlu 58
```

Db 306 ATTTGGCGGACGACGTGGGGCGCAACCCGACGGGAGCGCTTCGTGGCTGCAACGAG 365
Qy 59 CysGlyPheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsn 78
Db 366 TGGCCCTTCCCATCTCCGGGACTGTACAGTACGAGCGCGGAGGGCAGCGAAGC 425
Qy 79 CysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAsp 98
Db 426 TGGCCCCAGTGCAGAACCCGCTTCAAGCGCTCAAGGGGTGCGCGCGCTGCCGGGGAC 485
Qy 99 AspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGln 118
Db 486 GAGGAGAGGAGCGCGCTCGACACTCGAGAACAGATTCACTGGAGCGACAGCAGCAG 545
Qy 119 GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
Db 546 TCCGAGTACCTC-----GCCGAGTCCATGCTCCACGCC 578
Qy 139 ArgMetSerTyrGlyArgGlyProAsp---AspGlyAspGlyAsnAsnThrProGlnIle 157
Db 579 CACATGAGCTACGCGCGCGCGACCTCGACGCGGTGCGGACCATTCACCCCATC 638
Qy 158 ProProlIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsn 177
Db 639 CCC-----AATGTTCCCTC-----CTCACCAAC 662
Qy 178 GlyTyrGlyHisGlyGluValSerSerLeuHis----- 189
Db 663 GGACAGATGGTCGATGACATCCCGCGGACCGACCGCCCTTGTGCCCTCGTTCGTGGGT 722
Qy 190 -----LysArgIleHisProTyrProValSerGluProGly-----Ser 202
Db 723 GCGCGGGGAGAGGATTCACCCCTCCCGTACGCGGATCCCAACCTTCCTGTGCAACCG 782
Qy 203 AlalysTrpAspGluLysGlu-----ValSerTrpLys 214
Db 783 AGGTCATAGGACCTTCCAGGATCTCGCCCATATGGCTACGGGAGCGTAGCATGGAG 842
Qy 215 GluArgMetAspAspTrpLysSerLysGlnGlyIleLeu-----Gly 228
Db 843 GAGAGGATGGAGCGTGGAGCAGAGCAGGAGGAGTGCACACAGAGGAAAGCATGGC 902
Qy 229 GlyGlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGln 248
Db 903 GCGCGCGATGATGGTAT-----GATGCAATCTTACCACCTAATGGATGAAGCTAGACAG 956
Qy 249 ProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIle 268
Db 957 CCATTGTCAGAAAGATCCCGCTTCCTTCAAGCCCAATCAACCCCTATAGGATGATATA 1016
Qy 269 ValValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValPro 288
Db 1017 ATAATTCGGCTAGTGTGTTGTTGTTCTTCTTCCACTACCGAGTGCATCCCGTGCCT 1076
Qy 289 AspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrp 308
Db 1077 GATGCATTGCTTATGGCTCATATCTGTGATCTGTGAATTTGGTTGGCCATGTCCTTGG 1136
Qy 309 IleLeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeu 328
Db 1137 ATCTTGACCAAGTTCCAAAGTGGTTCTTATCGAGAGGAAACCTATCTTACCGCGCTG 1196
Qy 329 SerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheVal 348
Db 1197 AGTTAAAGTTTGACAAAGGAGGATCTCTCTCAACTCGCCCTGTGCAATTTCTTTGTC 1256
Qy 349 SerThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIle 368
Db 1257 AGTACGGTGTATCCCTTGAGGAACCTCATTTGGTACCTGCTTAATACCTTCTATCTATC 1316
Qy 369 LeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspGlyValaSer 388
Db 1316 TTTTGTGATCCACACTTCTTGAATGGTGGAAACCCCTGGAAGTGGCCAGCTTCT 2453

Db 1317 CTTTCGGTGGATTTATCAGTTGATAAGGTTTCATGCTAGCTTCTGATGATGGTGTGCC 1376
Qy 389 MetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPhe 408
Db 1377 ATGCTGACATTTGAAGCATTTGCTGAAACATCTGAATTTGCAAAAGAAATGGGTTCCTTC 1436
Qy 409 CysLysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAsp 428
Db 1437 TGCAAAAGATATAGCTTGGAGCTCTGCTCCAGAGTGGTACTTCCACAGAGATAGAC 1496
Qy 429 TyrLeuLysAspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGlu 448
Db 1497 TACCTGAAAGACAAAGTGGCGCCAAACTTTGTTAGAGAACGAGAGCAATGAAGAGAGAG 1556
Qy 449 TyrGluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAla 468
Db 1557 TATGAGGAATTCAGGTGCAATCAATGCTTGGTCTAAAGCCCAAAAGGTTCTTCGAG 1616
Qy 469 GluGlyTrpIleMetLysAspGlyThrProTyrProGlyAsnAsnThrArgAspHisPro 488
Db 1617 GAAGGATGGCAATGGAGATGGAACTCCATGCGCCGGAATTAATGTCCTGATCATCT 1676
Qy 489 GlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeu 508
Db 1677 GGAATGATTGAGTTCCTTGGTCAAAGTGGTGGCCATGATGTGGAAGGAATGAGCTG 1736
Qy 509 ProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAla 528
Db 1737 CCTCGATTTGTTTATGTTTCAAGAGAAAACGCGCCAGGCTACCAACCATCAAGAGGCT 1796
Qy 529 GlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeu 548
Db 1797 GGTGCTATGAATGATGTTGGTCCGAGTCTCTGCTGTACTAATAATGCTCTTATTTGCTG 1856
Qy 549 AsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPhe 568
Db 1857 AACTTGGATTGTGATCACTATATCAATAATAGTAAGCTATAAAGGAAGCAATGTGTTTT 1916
Qy 569 LeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAsp 588
Db 1917 ATGATGATCCTTTGCTTGGAAATAAAGTTTGTATGTGAGTTTCTTCAAAAGATTGAT 1976
Qy 589 GlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMet 608
Db 1977 GGGATTGATCCCATGATGATGCTTAAACAGAAATGTTGCTTTTCGATATCAACATG 2036
Qy 609 LysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArg 628
Db 2037 AAAGGTTTGGATGTATCCAGGGCCCAATTTATGTGGGTACTTGGATGTGCTTCCAGAGG 2096
Qy 629 GlnAlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCys 648
Db 2097 CAGGCATTATATGCTACGATGCTCCCAAA---ACAAGAGGCCCAACCATCAAGAACTTGC 2153
Qy 649 Asp-----CysCysProCysPheGly---ArgLysLysArgLys 660
Db 2154 AACTGCTGGCCAAAGTGGTGCATTTGCTGTGCTGTTTGGTAACAGAGAACCAAGAG 2213
Qy 661 HisAlaLysAspGlyLeuProGlu----- 668
Db 2214 AAGACCAAGACCTTAAACCTAAATTTGAGAAGATAAAGAAACTTTTTTAAGAAAAGGAA 2273
Qy 669 -----GlyThrAlaAspMetGlyValAspSer 677
Db 2274 AATCAAGCCCTGCTATGCTTTGGTGAATTTGATGAAGTGTCTCCAGGAGCTGAAAT 2333
Qy 678 AspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAla 697
Db 2334 GAAAGGCTAGTATTGTAATCAACAGAGTTGGAAAAGAAATTTGGCCAGTCTTCAGTT 2393
Qy 698 PheValThrSerThrLeuMetGluGlyValProProSerSerSerProAlaAla 717
Db 2394 TTTTGTGATCCACACTTCTTGAATGGTGGAAACCCCTGGAAGTGGCCAGCTTCT 2453

```
QY 718 LeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGly 737
Db 2454 CTTCTGAGGAAGCTATACATGTCATCAGCTGGATATGAAGACAAACAGCTGGGA 2513
QY 738 LeuGluLeuGlyTyrIleTyrGlySerIleThrGluAspIleThrGlyPheLysMet 757
Db 2514 AAGATATTTGGTTGGATTATGGATCAGTCACAGAGAATATTCTTACTGGGTTTAAGATG 2573
QY 758 HisCysArgGlyTyrArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySer 777
Db 2574 CACTGCCATGTTGGCGGTCAATTTACTGTCATACCTAAACGGCGCCTTCAAGGTTCC 2633
QY 778 AlaProIleAsnLeuSerAspArgLeuAenGlnValLeuArgTyrAlaLeuGlySerVal 797
Db 2634 GCACCTCTCAATCTTCCGATCGTCTTCACCAAGTTCTTCGGTGGGCTCTTGGTTCAATT 2693
QY 798 GluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLys 817
Db 2694 GAAATTTCTTCAGCAACCACTGCCCTCTCTGGTATGGGTAT---GGTGGGACTAAAG 2750
QY 818 TrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeu 837
Db 2751 TTCTCTGGAAGGTTTTCGTACATTAACTCCATCGTATACCTTGGACATCTATCCGGTCT 2810
QY 838 LeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSer 857
Db 2811 TTGGCTATTGACATTTGCTGCCATCTGCTTGTCTGACAGGAAATTTATCACGCCAGAG 2870
QY 858 IleSerThrPheAlaSerLeuPheIleAlaLeuPheMetSerIlePheAlaThrGly 877
Db 2871 CTTAACATGTTGCCAGCGCTGGTTCATGTACATTTTCATCTGATCTTTTGTCTACGAGC 2930
QY 878 IleLeuGluMetArgTrpSerGlyValSerIleGluTyrTrpArgAsnGluInPhe 897
Db 2931 ATCTCTGGAATCAGATGAGTGGTGTAGCATCGATGACTGTGGAGAAACGACGAGTTT 2990
QY 898 TrpValIleGlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLysVal 917
Db 2991 TGGGTCATTGGAGGGGCTCTTTCACATCTCTTGTCTGTTCCAGGACTCTCTCAAGTCT 3050
QY 918 LeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspLysAsp 937
Db 3051 ATAGCTGGGTGTAGACAGGCTTCACTGTGACATCAAG-----GGCGAGACGACGAG 3104
QY 938 GluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeu 957
Db 3105 GAGTTCTCAGAGCTGTACATTCAAATGGACGACCCCTTCTGATACCTCCGACACCCCTG 3164
QY 958 LeuIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyr 977
Db 3165 CTCCTACTGAACCTTCATTGGAGTGGTAGCTGGCATCTCCAATGCGATCAACAAACGATAT 3224
QY 978 GlnSerTrpGlyProLeuPheGlyLysLeuPheAlaPheTrpValIleValHisLeu 997
Db 3225 GAATCATGGGGCCCCCTGTTCCGGGAAGCTCTTCTTTGCAATTTGGTGTATGCTCATCTT 3284
QY 998 TyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValValIle 1017
Db 3285 TACCGTTCTCAAGGGTCTGTTGGAGGCGAGACAGACGCCAACGATGTCTATTGTC 3344
QY 1018 TrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIle 1037
Db 3345 TGGTCCATCTCTCGCTTCGATCTTCTCGCTGCTTGGGTCCGATCGACCCCGTTCCTT 3404
QY 1038 ValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3405 GCGAAGGATGATGGTCCCCCTGTTGGAGGATGTGTCTGGATTGC 3449
RESULT 3
LOCUS AY103655 3897 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays PC0096398 mRNA sequence.
```

```
ACCESSION AY103655
VERSION AY103655.1 GI:21206733
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3897)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitelett,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 3897)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
source
1. 3897
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635496"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 0 Length: 3897
Score: 4061.50 Matches: 756
Percent Similarity: 79.50% Conservative: 128
Best Local Similarity: 67.99% Mismatches: 141
Query Match: 71.91% Indels: 87
DB: 3 Gaps: 15

US-10-627-132-30 (1-1052) x AY103655 (1-3897)

QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
Db 156 ATGGAGGCGAGCGCGGCTGGTGGCGGCTCCCAACACCGAGCTCGTGTCTATC 215
QY 21 Arg-----GlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGlu 38
Db 216 CGCCGCGACGGCGATCCCGGCCGAGCGCGGGAGCAGACGGCGGCGTGTGCCAG 275
QY 39 IleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGlu 58
Db 276 ATTTGCGCGACGACGTCGCGCTTGCCTCCCGGGGACCCCTTCGTGGCGTCAACGAG 335
QY 59 CysGlyPheProValCysArgProCysTyrGluArgGluArgGluGlyThrGlnAsn 78
Db 336 TGCCTCTTCCCGCTTGCCTGGGACTGTCTACGAATACGAGCGCGCGGAGCGGACG 395
QY 79 CysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAsp 98
Db 396 TGCCCCCAGTGCAGACTCTCGATACAGCGCTCAAGGGCTGCCAACGCTGTACCGGTAC 455
QY 99 AspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGln 118
```

Db	456	GAAGAGGAGCAGCGCGTTCGATGACCTTGGACAACAGAGTTCAACTCTGGACGCGCCATGACTCG	515
Qy	119	GLnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly	138
Db	516	CAG-----TCTGTGGCCGAGTCCAGTCTCTACGGC	545
Qy	139	ArgMetSerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIlePro	158
Db	546	CACATGAGCTACGGCCGTGA-----GGTGACCTAATGGCGGCCACAAAGCTTTC	596
Qy	159	ProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGly	178
Db	597	CAGCTC-----AACCCCAATGTTCCACTC-----CTCACCAACGGG	632
Qy	179	TyrGlyHisGlyGluValSerSerSerLeuHis-----	189
Db	633	CAAAATGTTGATGACATCCCAACCGGACGACGCGCTGGTGTCTTCTTCATGGGTGGT	692
Qy	190	-----LysArgIleHisProTyrProValSerGluProGly-----SerAla	203
Db	693	GGGGGAAGAGGATACATCCCTCTTATGCGGATCCAGCTTACCTGTGCAACCCAGG	752
Qy	204	LysTrpAspGluLysLysGlu-----ValSerTrpLysGlu	215
Db	753	TCTATGACCCATCCAAAGGATCTGTGTCATATGGGTATGGTATGTGTTCTTGGAAAGGAA	812
Qy	216	ArgMetAspAspTrpLysSerLysGln-----GlyIleLeuGlyGly	229
Db	813	CGGATGAGAAATGGAAAGCAGACAGAGAGGATGACACAGCGGGGAATGATGTGGT	872
Qy	230	GlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnPro	249
Db	873	GGTGATGATGGTGAC-----GATGCTGATCTACCACTAATGATGATGAGCAAGCAACAA	926
Qy	250	LeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleVal	269
Db	927	CTGTCCAGGAAAATTCACATCCATCAAGCCAGATTAAATCATATAGTAGTATCATTT	986
Qy	270	ValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAsp	289
Db	987	ATTCGGCTTTGGTTTTGGGGTCTCTTCCACTACCGAGTGATGCAATCCGGTGAATGAT	1046
Qy	290	AlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIle	309
Db	1047	GCATTTGCTTTGTGGCGCATATCTGTTATCTGTGAATCTGGTTGCGCATGCTTGGATT	1106
Qy	310	LeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSer	329
Db	1107	CTTGATCAATTCCAAAGTGGTTCCTTATTGAGAGAGACTTACCTAGACCGCGCTGTCA	1166
Qy	330	LeuArgTyrGluArgGlyGluProSerLeuLeuSerAlaValAspLeuPheValSer	349
Db	1167	CTGAGTTTCGAAGGAGGCGCCCACTCTCAACTTGTCTCCAAATGATTTCTTTGTTCAGT	1226
Qy	350	ThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeu	369
Db	1227	ACGGTTGATTCCTTAAGGAACCTTCCTTTGGTTCACAAATACTGTTCTATCTATCTT	1286
Qy	370	AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet	389
Db	1287	TCGGTGGATATTCCTGTTGATAAGGTTCTTGTCTATGTTTCTGATGATGGTGTGCAATG	1346
Qy	390	LeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCys	409
Db	1347	CTAACGTTTGAAGCATATATCTGAAACATCTGAATTTGCAAAAGTGGGTTCCTTCTGTC	1406
Qy	410	LysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyr	429
Db	1407	AAACGGTACAATATTGAACCTTCGGCTCCAGAGTGGTACTTCCAAACAGAGATAGACTAC	1466
Qy	430	LeuLysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArgGluTyr	449

1467	TTGAAAGACAAGGTGGCCACAACTTTGTTAGGGAGAGAGCAATGAAGAGAGNGTAT	1526
450	GIUGLUPHELysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProLaGIu	469
1527	GAGGAATTCAGGTGAGATCAATGCCTTAGTTGGCCAAAGCCAGAAAGTTCTGTGAAGA	1586
470	GLYTrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly	489
1587	GGATGACCAATGCAAGATGGAAACCCCTCGCTCGGCAACAATGTTCTGTGATCATCTCGA	1646
490	MetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuPro	509
1647	ATGATTCAGGTCTTCCTTGGCCAAAGCGAGGCTTGACTGTGAGGAAATGAATGCCCA	1706
510	ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisIshLysLysAlaGly	529
1707	CGATTGGTTTATGTTCTTAGAGAGAAACGACCAAGGCTATAACCATCATAAAGAAAGCTGT	1766
530	AlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsn	549
1767	GCTATGAATGCATTGGTCGAGTCTCTGCTGTACTAACCAATAGCTCCATATTTGTTAAAC	1826
550	LeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeu	569
1827	TTGGATTGTGATCACATACATCAACACAGCAAGGCTATAAAGAGACAAATGTTTATG	1886
570	MetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGly	589
1887	ATGGACCCCTTACTAGGAAGAAGGTTTGTCTATGTATACAGTTCCTCTCAAAGATTTGATGG	1946
590	IleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLys	609
1947	ATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTCATATCAACATGAAA	2006
610	GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGln	629
2007	GGTTTGGATGGTATTGAGGTCCAAATATATTTGGTACTGGATGTGTATTTAGAGGCAG	2066
630	AlaLeuTyrGlyTyrAsnProProLysArgProLysArgProLysMetValThrCysAsp	649
2067	GCATTATATGGTTATGATGCCCCCAA--ACAAGAAGCCACCATCAAGGACTTCGAAC	2123
650	-----CysCysProCysPheGlyArgLysLysArgLysHisAla	662
2124	TGCTGGCCCAAGTGTGCTTTTCTGTCTGCTTTGGCAATAGGAAGCAAAAGAGACT	2183
663	LysAspGlyLeuProGlu-----	668
2184	ACCAAAACCCAAAACAGAGAGAAAAGTATTTATTTTTCAGAAAGAGAGAACCAATCC	2243
669	-----GlyThrAlaAspMetGlyValAspSerAspLysGlu	680
2244	CTGCAATATGCTCTTGGTGAAATGACGAAGCTGCTCCAGGAGCTGAGATGAAAGGCC	2303
681	MetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaPheValThr	700
2304	GGTATTGTAATAACACAAAAATTAGAAAAAATTTGGCCAAATCTCTGTTTGTGTACA	2363
701	SerThrLeuMetGluGluGlyValProProSerSerSerProAlaAlaLeuLeuLys	720
2364	TCCACATCTTCGAGAAATGGTGAACCTTGAAGAGTGAAGTCTCTCTTTTGGAAA	2423
721	GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeu	740
2424	GAAGCTATATCATGTATTGTTGTGTTATGAGACACAGACAGCTGGGGAAGAAGATT	2483
741	GlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg	760
2484	GGCTGGATCTATGGATCAGTTACAGAAATATCTTAATCGTTTCAAGATGCAATTTGTCAT	2543
761	GlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIle	780
2544	GGTTGGCGGTCAATTTACTGCATCTCTTAAACGGGTGCATTTCAAAGGTTTGCACCTCT	2603

QY 781 AsnLeuSerAspArgLeuAenGlnValLeuArgTrpAlaLeuGlySerValGluLeuPhe 800
Db 2604 AATCTTTCAGATCGTCTTACCAGGTGCTTGGGTCTTGGGTCTTATGAGATCTTC 2663
QY 801 PheSerArgHisProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGlu 820
Db 2664 TTCAGCAATCATGCTCTTTCGTAAGGTAT---GGTGGCGGTCTGAATATTTTGGAA 2720
QY 821 ArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyr 840
Db 2721 AGATTTTCTACATCACTCCATCGTATCTTGGACATCTATTTCCCTCTTGGCTTAC 2780
QY 841 CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr 860
Db 2781 TGTACATTCCTGCCATCTGTTATTGACAGGAAATTTATCACTCCAGAGCTGAATAAT 2840
QY 861 PheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyLeuGlu 880
Db 2841 GTTCCAGAGCTGTGGTTCATGTCACCTTTTATCTGCAATTTTGTACGAGCATCTAGAA 2900
QY 881 MetArgTrpSerGlyValSerIleGluGluTyrTrpArgAsnGluGlnPheTrpValIle 900
Db 2901 ATGAGATGGAGTGGTTCGATTCGATCTGTTGGAGGAATGAGCATCTTGGGTCAAT 2960
QY 901 GlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLysValLeuAlaGly 920
Db 2961 CGAGGTGTCTCTCACCTCTTGTCTGTGTTCCAGGACTTCTCAAGGTCTAGCTGT 3020
QY 921 IleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPheAla 940
Db 3021 GATGATACAGCTTCCAGGTGATCAAAAG-----GGTGGAGATGATGAGAGTCTCA 3074
QY 941 GluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeuLeuIleIle 960
Db 3075 GAGCTATATACATTCAAATGAGTACCTTATTGATACCTCTTACCACCTTGTCTTATG 3134
QY 961 AsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTrp 980
Db 3135 AACTTCATTTGGTGTCTGCTGGCTTCAAAATGCGATCAATAACGATATGAGTCATGG 3194
QY 981 GlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPhe 1000
Db 3195 GCGCCCTCTTTGGGAAGCTATCTTGTCAATTTGGGTGATTTGTCATCTTTATCCCTTT 3254
QY 1001 LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValValIleTrpSerIle 1020
Db 3255 CTCAAAGGTTTGGTTGGAAGGCAAAACAGACACCAAGATTTGTCATCTGCTGCTCAAT 3314
QY 1021 LeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThr 1040
Db 3315 CTGCTGGCTTCAATCTTCTCGCTCTTGGGTTCGATTTGATCTTCTTCTCGGAAGAT 3374
QY 1041 LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3375 GATGTCCTGCTTCTTGGAGAGTGTGGTTGGATTGC 3410
RESULT 4
AY103701 Zea mays P020363 mRNA sequence. HTC 16-OCT-2002
LOCUS
DEFINITION Zea mays P020363 mRNA sequence.
ACCESSION AY103701
VERSION AY103701.1 GI:21206779
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD.
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 3788)
REFERENCE
AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of

Overgo Probes
Unpublished (2002)
2 (bases 1 to 3788)
Coe, E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB. www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI. www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schmable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
Location/Qualifiers
1..3788
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:637005"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
Alignment Scores:
Pred. No.: 0 Length: 3788
Score: 3938.00 Matches: 751
Percent Similarity: 78.26% Conservative: 120
Best Local Similarity: 67.48% Mismatches: 161
Query Match: 69.72% Indels: 81
Gaps: 17
DB: 3
US-10-627-132-30 (1-1052) x AY103701 (1-3788)
QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
Db 201 ATGAGGCTAGCGGGGCTGGTGGCGGCTCGCATACCGGAACGAGCTGTGTGTATC 260
QY 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnVal----- 36
Db 261 CGC-----CGCAGCCGAGTCCGGAGCCCGGGCGGGCGGGCGGGCGGAG 314
QY 37 -----CysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheVal 54
Db 315 GCGCGGTGCCAGATATGCGGACGAGTCCGGGTGGGCTTCGACGGGAGCCCTTCGTG 374
QY 55 AlaCysAsnGluCysGlyPheProValCysArgProCysTyrGlyTyrGluArgGlu 74
Db 375 GCGTGCACAGTGCCTTCCCGTCTGCGCGCTGTCTACGAGTACGAGCGCGCGAG 434
QY 75 GlyThrGlnAsnCysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArg 94
Db 435 GGCTCGCAGGCGTCCCGCAGTGCAGGACCCGCTACAGCGCTCAAGGGTTCGCCGCG 494
QY 95 ValAlaGlyAspAspGluGluAspIleAspLeuGluHisGluPheAsnIleAsp 114
Db 495 GTGGCTGGGACGAGGAGGAGCGGCTCGACGCTGGAGGGGAGTTCGGCCTGCGAG 554
QY 115 AspGluAsnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGln---IleThrGlu 133
Db 555 GACGGC-----GCCGGCCACGAGGACGCCCGCAGTACGTCCGCCAG 596
QY 134 AlaMetLeuHisGlyArgMetSerTyrGlyArgGlyProAsp---AspGlyAspGlyAsn 152
Db 597 TCCATGTGCGGGCGCAGATGATGATCGGCGCGGGCGGCGGCGGAGCCCGAGCTTCAAC 656
QY 153 AsnThrProGlnIleProProIleIleThrGlySerArgSerValProValSerGlyGlu 172

Db	657	CCCGTCCCAACGTCGCGCTCTCCACCAACGGCCAGATGGTTGATCACTCCGCGGAG	716
Qy	173	-----PheProIleThrAsnGlyTyrGlyHisGlyGluValSerSerLeu	188
Db	717	CAGCACGGCTCGTGGCTCTCATATGGGCGGCGCGCGC	758
Qy	189	HisLysArgIleHisProTyrProValSerGluProGly-----SerAlaLys	204
Db	759	GGCAGAGGATCCACCGCTCCCTTCGCTGATTCCAACTTCAGGTGCAACCGAGATCC	818
Qy	205	TrpAspGluLysLysGlu-----ValSerTrpLysGluArg	216
Db	819	ATGACCCGCTCAAGGATCTGGCGCGCTACGGATATGGCAGCGTGGCTCGAAGGAGAGA	878
Qy	217	MetAspAspTrpLysSerLysGlnGlyLeu-----GlyGlyGlyAla	231
Db	879	ATGAGGGCTGGAAAGACAGACAGCGCTCGAGCATGTCTAGGACGAGGGTGGCGGT	938
Qy	232	AspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnProLeuSer	251
Db	939	GATTGGGATGGCAGCATGCAGATCTGCCACTAATGGATGAAGCTACGACGCCATTGTC	998
Qy	252	ArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleValValArg	271
Db	999	AGAAAAGTCCCTATATCATCAAGCCGAATAATCCCTACAGGATGATTATCGTTATCCGG	1058
Qy	272	LeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAspAlaIle	291
Db	1059	TTGTGGTTTTGGTTTTCTCTCCACTACCGAGTGTGATCATCCGCGAAGATGCATTT	1118
Qy	292	GlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAsp	311
Db	1119	GCATTGGCTCATATCTGTAATCTGTGAATCTGGTTTGGCATGTCTTGGATCTTGAT	1178
Qy	312	GlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSerLeuArg	331
Db	1179	CAGTTCCAAAGTGGCTTCCAATCGAGAGAGAGACTTACCTGGACCGTTGTCTCACTAAG	1238
Qy	332	TyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSerThrVal	351
Db	1239	TTTGACAAAGGAAGTCAACCCCTCTCAGCTGTCTCCAATCGACTTCTTTGTGACGACGGT	1298
Qy	352	AspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAlaVal	371
Db	1299	GATCCCAAGAAAGAACCTCCCTTGGTCACAGCAACACTGTCTCTTTTCCATCTTCTGTG	1358
Qy	372	AspTyrProValAspLysValSerCysTyrValSerAspGlyAlaSerMetLeuThr	391
Db	1359	GATTATCCGGTTGAGAAGTCTCCTGCTATGTTTCTGATGATGGTGTGCAATCTTACG	1418
Qy	392	PheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLysLys	411
Db	1419	TTTGAAGCATTTGTCTGAACACTGAAATTCGAAGAAATGGGTTCCTTTTCAGCAAAAG	1478
Qy	412	PheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeuLys	431
Db	1479	TTTTAATATCGAGCTCTGCTGCTCTGTGGTGTACTTCCAAAGAGATAGACTACTGAA	1538
Qy	432	AspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGluTyrGluGlu	451
Db	1539	GACAAAGTTGTGCTTCTTATTTAGGAGAGAGGCGCGCATGAAGAGAGAAATACGAGAA	1598
Qy	452	PheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGluGlyTrp	471
Db	1599	TTCAGGTAAAGATCAATGCTCTGGCTGGAAACAACTGACCGCATCATCTTGGATGATT	1658
Qy	472	IleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGlyMetIle	491
Db	1659	ACATGCAAGATGGAAAGCCCTGGCTGGAAACAACTGACCGCATCATCTTGGATGATT	1718
Qy	492	GlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeu	511

1719	Db	CAGGTATTCCTTGGCCAAAGTGGCGGTCTGTGATGTGGAAGGAAATGACGTGTGCTCGCCTG	1770
512	Qy	ValTyrValSerArgGluLysArgProGlyPheGlnHisHisGlyLysAlaGlyAlaMet	531
1779	Db	GTATTATGTCGAGAGAAAGAGCCAGGTATTAACCATCACAAAGAGGCTGGTGCCATG	1838
532	Qy	AsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAsp	551
1839	Db	AATGCACCTGGTCGTCTCTGTCTTATCAAAATGCTGCATACCATATTGAACTTGGAC	1898
552	Qy	CysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeuMetAsp	571
1899	Db	TGTGATCACTACATCAACAATAGCAAGGCCATAAAGAGCGTATGTGTTTCATGATGGAT	1958
572	Qy	ProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIleAsp	591
1959	Db	CTTTTGGTGGGNAAGAAAGTGCTATGTATACAGTTCCCTCAGAGGTTTGATGTATTGAC	2018
592	Qy	ValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLysGlyLeu	611
2019	Db	AAAAATGATCGATACGCTAACAGGAACGTTGTCTTTTTCACATCAACATGAAGAGTTTG	2078
612	Qy	AspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeu	631
2079	Db	GACGGTATTCAAGGACCCATTATGTGGGTACTTGGATGTGTTTTTCAGACGGCAGGCACTG	2138
632	Qy	TyrGlyTyrAsnProLysGlyProLysArgProLysMetValThrCysAspCysCys	651
2139	Db	TATGGTTATGATGCTCTTAA---ACGAGAAGCCCATCAAGAACTTGCAACTGCTGG	2195
652	Qy	Pro-----CysPheGlyArgGlyLysArgLys-----	660
2196	Db	CCCAAGTGGTGCCTCTCTTCTGCTGCAGCAGCAAGAAATAAAAGAAAGACTACAAAA	2255
660	Qy	-----	660
2256	Db	CCAAAGCAGGAGAGAAAGAGATTATTTTCAAGAACGACAGAAACCCTCTCTCTGCA	2315
661	Qy	HisAlaLysAspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAspLysGlu	680
2316	Db	TATGCTTTGGTGAAATGTATGATGAAGGTGCTCCA-----GGTCTGATATCGAGAAGGCC	2369
681	Qy	MetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThr	700
2370	Db	GGAATCGTAATCAACAGAAACTAGAGAGAAATTTGGGCAGCTTCTCTGTTTGTCTGCGA	2429
701	Qy	SerThrLeuMetGluGluGlyGlyValProProSerSerSerProAlaAlaLeuLys	720
2430	Db	TCAACACTCTTGAGAACGGAGGAGCCCTGAAGAGCGCAAGTCAGCTTCTCTCTGAG	2489
721	Qy	GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeu	740
2490	Db	GAAGCTATACATGTTATCAGCTGGGCTACGAAGACAAGACCGNACTGGGNAAGAGATT	2549
741	Qy	GlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg	760
2550	Db	GGCTGGATTATTCGGATCGATCGACAGAGGATATCTTGATCGATTAAAGATGCACTGCCAT	2609
761	Qy	GlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIle	780
2610	Db	GGCTGGCGGTCTATTTTACTGCATCCCGAAGCGGCTGCATTCAAAGGTTCTGCCCTCTG	2659
781	Qy	AsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePhe	800
2670	Db	AACCTTTCCGACCGTCTTCCACGAGTCTTTCGCTGGGCGCTTGGTCCGTCGAAATTTTC	2729
801	Qy	PheSerArgHisSerProLeuLeu-TyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGlu	820
2730	Db	TTTCAGNAANCAATGNCACCTTTGGTACGGATAC---GGCGCGCGGCTGAAATTCCTGGA	2786
820	Qy	uArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyr	840
2787	Db	AAGTTTTCTTATATCAACTCCATCGTTTATPCCCTGGACGTCATCTCTCTCTGGCTTA	2846

Db 1007 GGCACAAGCATGTCCTCCCTCTGAGGGTCGGGGTGTGGTGATATTGATGCATCAACTGAT 1066
 Qy 239 -----AapValAlaLeuAsnAapGluAlaArgGlnProLeuSerArgLysVal 254
 Db 1067 TACAACATGGAAGATGCCCTATTGAACAGCAAGAACTCGACAGCCCTCTATCTAGGAAAGTT 1126
 Qy 255 SerileAlaSerSerLysValAsnProTyrArgMetValIleValValArgLeuValVal 274
 Db 1127 CCATCTCTCTCCATGAGTAAATCCATACAGGATGGTCATGTGCTGCGGATTAATGTT 1186
 Qy 275 LeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAapAlaIleGlyLeuTyr 294
 Db 1187 CTAGCATCTCTTGCATACCGTATCACAAATCCTGTGCGCAATGCATACCCATTATGG 1246
 Qy 295 LeuValSerIleLeuCysGluIleTyrPheAlaIleSerTyrIleLeuAapGlnPhePro 314
 Db 1247 CTTCTATCTGTATATGTGAGATCTGGTTTCTCTCTTCTGGATATTTGGATTCAGTTCCT 1306
 Qy 315 LysTyrPheProIleAapArgGluThrTyrLeuAapArgLeuSerLeuAatGlyTyrGluArg 334
 Db 1307 AAGTGGTITTCATCAATCAACGGGAGACGTACCTTGTATAGGCTGGCATTAAGGTATGACCGG 1366
 Qy 335 GluGlyGluProSerLeuLeuSerAlaValAapLeuPheValSerThrValAapProLeu 354
 Db 1367 GAAGGTGAGCCATCTCAGTTGGCTGCTGTGCATTTTCGTACATTTTCGTACGTGACCCCAATG 1426
 Qy 355 LysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAlaValAapTyrPro 374
 Db 1427 AAGGAGCCTCTCTGTGTCACTGCCAATACCGTGTATCCATTTCTGTGTGGATTACCCCT 1486
 Qy 375 ValAapLysValSerCysTyrValSerAapGlyAlaSerMetLeuThrPheGluSer 394
 Db 1487 GTGGATAAGGTCTCTGTGTATGTATCTGATGAGTGGCATGTGCATTTTGAATGCA 1546
 Qy 395 LeuSerGluThrAlaGluPheAlaArgLysTyrValProPheCysLysLysPheGlyIle 414
 Db 1547 CTAGCTGAGACTTCAGAGTTTGTCTAGAAAATGGGTACCATTTGTTAAGAAGTACAACTT 1606
 Qy 415 GluProArgAlaProGluPheTyrPheSerLeuLysValAapTyrLeuLysAapLysVal 434
 Db 1607 GAACCTAGAGCTCTCTGAATGGTACTCTCCAGAAAATTTGATTACTTGAAGGACAAAGTG 1666
 Qy 435 GlnProThrPheValGlnGluArgAlaMetLysArgGluTyrGluGluPheLysVal 454
 Db 1667 CACCTCTCATTTGTTAAGACCGCGGCCCATGAAGAGATATGAAGAATTCAAAGTT 1726
 Qy 455 ArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGluGlyTyrIleMetLys 474
 Db 1727 AGGGTAAATGGCTTTGTTGAAGGCACAGAAAAGTTCTCTGAGGAAGGATGCATGCAA 1786
 Qy 475 AspGlyThrProTyrProGlyAsnAsnThrArgAapHisProGlyMetIleGlnValPhe 494
 Db 1787 GATGGCACACCATGGCCAGGAACAATAACCMGGGACCATCTCTGGAATGATTGAGTTTC 1846
 Qy 495 LeuGlyHisSerGlyGlyHisAapThrGluGlyAsnGluLeuProArgLeuValTyrVal 514
 Db 1847 CTGTGTACAGTGGTGGCTTGATCTAGGGCAATGAGCTACCCCGCTTGTGTATGTT 1906
 Qy 515 SerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeu 534
 Db 1907 TCTCGTGAAGAAGCGTCTGGATTCAGCATCACAGAAAGCTGGTGCCATGAATCTCTT 1966
 Qy 535 IleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAapCysAapHis 554
 Db 1967 GTTCGTGTCTCAGCTGTGCTTACCAATGGACAATACATGTTGAAATCTTGAATGTGATCAC 2026
 Qy 555 TyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPhePheLeuMetAapProGlnVal 574
 Db 2027 TACATTAACACAGTAAGGCTCTCAGGGAAGCTATGTGCTTCTTATGGACCTTAACCTA 2086
 Qy 575 GlyArgLysValCysTyrValGlnPheProGlnArgPheAapGlyIleAapValHisAap 594

Db 2087 GGAAGGAGTGTCTGTCTACGTCCAGTTTCCCCAGAGATTCCGATGGCAATGAT 2146
 Qy 595 AatGlyTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLysGlyLeuAapGlyIle 614
 Db 2147 CGATATGCCAACAGGAACACCGTGTGTTTTCGATATTAACTTGAGAGTCTTGATGGCATC 2206
 Qy 615 GlnGlyProValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyr 634
 Db 2207 CAAGGACCACTTATGTCGGAACCTGCTGTGTTTCAACCGAACAGCTCTATATGGTTAT 2266
 Qy 635 AsnProProLysGlyProLysArgProLysMetValThrCysAapCysCysProCysPhe 654
 Db 2267 GAGCCCCCAATTAAGCAGAGAAGAGGGTGTGTTCTTGTCA-----TCACATATGGGC 2317
 Qy 655 GlyArgLysLys-----ArgLysHis 661
 Db 2318 GGTAGGAAGAAGGCAAGCAAAATCAAGAAGGGCTCGCAAGAAGAAGTCGCAGAGCAT 2377
 Qy 662 AlaLysAapGlyLeuPro-----GluGlyThrAlaAapMet 673
 Db 2378 GTGCACAGTTCTGTGCCAGTATTCAACCTTTGAAGATATAGAGGAGGAGTTGAAGGCGCT 2437
 Qy 674 GlyValAapSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGly 693
 Db 2438 GGATTTGACGACGAGAAATCACTTCTTATGCTCTCAATGAGCCTGGAGAAGAGATTTGGC 2497
 Qy 694 GlnSerAlaAlaPheValThrSerThrLeuMetGluGluGlyValProProSerSer 713
 Db 2498 CAGTCCGACAGGTTGTGTGCTCCACTCTGATGAGTATGTTGTTCTCTCAGTCGCA 2557
 Qy 714 SerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAapLys 733
 Db 2558 ACTCGGAGTCTCTCTGNAAGAAGCTATCCATGTTATAAGCTGTGCTATGAGCAAG 2617
 Qy 734 ThrAapTyrGlyLeuGluLeuGlyTyrIleTyrGlySerIleThrGluAapIleLeuThr 753
 Db 2618 ACTGAATGGGAACTGAGATCGGTGGATCTACGGTTCTGTGACAGAAGACATTTCTCACC 2677
 Qy 754 GlyPheLysMetHisCysArgGlyTyrArgSerValTyrCysMetProLysArgAlaAla 773
 Db 2678 GGATTTCAAGATGACCGCGGAGGCTGGCGTGCATCTACTGTGATGCCCAAGGCGCCAGCT 2737
 Qy 774 PheLysGlySerAlaProIleAsnLeuSerAapArgLeuAsnGlnValLeuArgTyrAla 793
 Db 2738 TTCAAGGGGTCTGCCCCCATCAATCTTTCGACCGCTCTGAACACAGGTGCTCCGTGGCT 2797
 Qy 794 LeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsn 813
 Db 2798 CTTGGGTCCGTGGAGATCTCTTTTCAGCGGCACCTGCCCCCTGTGGTACGGGTAC---GGA 2854
 Qy 814 GlyAsnLeuLysTyrLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThr 833
 Db 2855 GGGCGGTCTCAAGTTCTCTGGAGAGATTCGGTACATCAACACCATCTACCCGCTCAGC 2914
 Qy 834 SerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPhe 853
 Db 2915 TCCATCCGCTTCTCATCTACTGTCATCTGCGCCCATCTGTCTGTCTCACCGGAAAGTTC 2974
 Qy 854 IleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIle 873
 Db 2975 ATCATTCAGAGATCAGCACTTCGCCAGCATCTGGTTTCATCTCCCTCTTCATCTCGATC 3034
 Qy 874 PheAlaThrGlyIleLeuGluMetArgTyrSerGlyValSerIleGluGluTyrTyrArg 893
 Db 3035 TTCGCCACGGCATCTCTGGAGATGAGTGGAGCGGGGTGGGCATCGACAGATGGTGGAGG 3094
 Qy 894 AsnGluGlnPheTyrValIleGlyLysValSerAlaHisLeuPheAlaValIleGlnGly 913
 Db 3095 AACGAGCAGTTCTGGGTGATCGGGGCATCTCCGCGCACCTCTTCGCGCTGTTCAGGGC 3154
 Qy 914 LeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGly 933
 Db 3155 CTGTCAAGGTGCTGGCGGCATCGACACCAACTTCACCGTCACCTCCAGGCGCTCG--- 3211


```

QY      934  AspGluAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuLeuPro 953
Db      3212 GACGAGGACGGGACCTTCGCGAGCTGTACATGTTCAAGTGGACGACGCTCTGTATCCCG 3271
QY      954  ProThrThrLeuLeuLeuLeuValIleGlyValValAlaGlyLeuSerAspAlaIle 973
Db      3272 CCACACCATCTGATCATCAACCTGGTGGCGTCTGCGCGGATCTTCTTACGCCATC 3331
QY      974  AsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpVal 993
Db      3332 AACAGCGGATACAGTCGTCGGGCGCTCTTCGGCAAGCTTCTTCTGCTTCTGGGTC 3391
QY      994  IleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThr 1013
Db      3392 ATCGTCCACCTGTACCGCTTCTCAAGGCGCTCATGGGCGAGCAGACCGCACCGGACC 3451
QY      1014 ValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIle 1033
Db      3452 ATCGTCGTCTGGGCCATCTGTCGGCGTCCATCTTCTCTGCTGTGGGTTCGCATC 3511
QY      1034 AspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db      3512 GACCCCTTCACCAACCGGCTCACTGGCCCGATACCCAGACGCTGGCATCAACTGC 3568

RESULT 6
LOCUS   CL972423
DEFINITION Oryza sativa (indica cultivar-group)
ACCESSION CL972423.1
VERSION   CL972423.1
KEYWORDS  GSS.
SOURCE   Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
REFERENCE 1 (bases 1 to 3192)
AUTHORS   Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
           Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
           Wong,G.K.S., Deng,X.W. and Wang,J.
           An analysis of transcriptional regulation of the rice genome and
           its comparison to Arabidopsis
           Unpublished (2004)
           Contact: Chen Chen
           Department of Bioinformatic
           Beijing Institute of Genomics
           Chinese Academy of Sciences, Beijing 101300, China
           Tel: 86-10-80481559
           Fax: 86-10-80488676
           Email: chenchen@genomics.org.cn
           Rice genomic sequence.
           Class: exon-trapped.
FEATURES             Location/Qualifiers
     source           1..3192
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:39946"
                     /clone_lib="Oryza sativa Exress Library"
                     /note="Oryza sativa exon trapped genomic sequences"

ORIGIN
Alignment Scores:
Pred. No.: 0
Score: 3907.00
Percent Similarity: 79.42%
Best Local Similarity: 70.11%
Query Match: 69.17%
DB: 9
Length: 3192
Matches: 753
Conservative: 100
Mismatches: 144
Indels: 78
Gaps: 14

US-10-627-132-30 (1-1052) x CL972423 (1-3192)

```

```

QY      33  SerGlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeu 52
Db      43  AGCGGCGAGGGGTGCGAGATCTCGCGCGACGGCGTGGGCACGACGCGCGGCGGCGACGCG 102
QY      53  PheValAlaCysAsnGluCysGlyPheProValCysArgProCysTrpGlyTrpGluArg 72
Db      103  TTGCGCGCTGGAGCGTCTGCGGCTTCCGGGTGTCGCCGCCCTGCTAGAGTACGAGCGC 162
QY      73  ArgGluGlyThrGlnAsnCysProGlnCysValysThrArgTyrLysArgLeuLysGlySer 92
Db      163  AAGGATGCGACCCAGCGCTTGGCCCCAGTGCAGACCAAGTACAGGCGCCACAGGGGAGC 222
QY      93  ProArgValAlaGlyAspAspGluGlu---AspIleAspAspLeu---GluHisGlu 110
Db      223  CCGCGCATCTCGTGGGAGGAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGAT 282
QY      111  PheAsnIleAspAspGluAsnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGln 130
Db      283  TACCCTGTCATCTGGCAGTGGCGACCGAAGCAG-----AAG 318
QY      131  IleThrGluAlaMetLeuHisGlyArgMetSerTyrGlyArgGlyProAspAspGlyAsp 150
Db      319  ATTGCTGATAGGATGCGCAGTGGCGCATGAATGCTGGGGTGGTGGAGATGCTCGGCCGT 378
QY      151  GlyAsnAsnThrProGlnIleProProIleIleThrGly--SerArgSerValProVal 170
Db      379  -----CCCAAGTATGACAGTGGCGAGATCGGCTCACAATACT 417
QY      170  erGlyGluPheProIleThrAsnGlyTyrGlyHisGlyGluValSerSerSerLeuHisL 190
Db      418  CGGAGAAATCCCTCGGAGCTTCCCTGATCATCATATGATGATGATGATGATGATGATGATGAT 477
QY      190  ysArgIleHisProTyrProValSerGluProGlySerAlaLys-----T 205
Db      478  GGCAAGCGTGTCCATTCCTCCCTATGTGAACCATTCACCAATCCATCAAGGAGTCTCT 537
QY      205  rpAspGluLysGluValSerTrpLysGluArgMetAspAspTrpLysSerLysGlnG 225
Db      538  GGCAGCATTTGAAAT--GTTGCTGGAAAGAAAGAGTTGATGGCTGGAAACTGAAGCAG- 594
QY      225  lyIleLeuGlyGlyGlyAlaAspPro----- 233
Db      595  -----GACAAGGGAGCAATTCCTCATGCAATGGGACAGCATTCGCCCTTCTGAG 646
QY      234  -----GluAspMetAspAla-----AspValAlaLeuA 243
Db      647  GTCGGGAGTGGTGTATCGATCGATCGATCCACTGATTACAATATGGAAGATGCCTTACTGA 706
QY      243  snAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSerSerLysValAsn 263
Db      707  ATGATGAAACTCGCGCAGCTCTCTCTAGAAAAGTTCCCTTCTCATCCAGAAATAATC 766
QY      263  roTyrArgMetValIleValValArgLeuValLeuAlaPhePheLeuArgTyrArgI 283
Db      767  CTTACAGATGGTCAATGTTCTGCTGGTGGTGTCTTAAGCATTTTCTTCTCATCCCGTA 826
QY      283  leLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerIleCysGluIleT 303
Db      827  TTACGAATCTGTGGCAATCGGTATCCGCTCTCGCTTTTATCTGTTATATGTGAGATT 886
QY      303  rpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPheProIleAspArgGluT 323
Db      887  GGTTCGCTTCTGCGATATTGGATCAGTTCGCGAGTGGTTCCTCAATCAACCGTGA 946
QY      323  hrTyrLeuAspArgSerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerA 343
Db      947  CCTACCTGTAGTGGCTGGCATTTGAGGTATGACAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1006
QY      343  laValAspLeuPheValSerThrValAspProLeuLysGluProProLeuValThrAlaA 363
Db      1007  CTGTTGACATTTTGTTCAGTACAGTCCGACCCCATGAGAGAACCTCTCTTGTACTGCA 1066
QY      363  snThrValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrVal 383

```

```
Db 1067 ATACCGTGCTGCTCAATCTTCGTGCTGATTAACCCAGTGGACAGGTCTCTTCGTATGAT 1126
Qy 383 erAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaA 403
Db 1127 CTGATGATGGTGTGCTGCAATGCTGACTTTTGTATGATGATTTGGCTGAGACTTTCAGAGTTTGCTA 1186
Qy 403 rglYsTrpValProPheCysAlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGluPheValP 423
Db 1187 GAAAGTGGGTTCCTATTCGTTAAAGAGTATTAACATTGAGCCCAAGAGCTCCTGAGTGGTACT 1246
Qy 423 heSerLeuLysValAspTyrLeuLysAspLysValGlnProThrPheValGlnGluArgA 443
Db 1247 TCTCCAGAAATTCGATTCTTGAAGACAAAGTCCACCCCTTCATTGTTGTTAAAGACCGCC 1306
Qy 443 rgaLaMetLysArgGluTyrGluGluPheLysValArgLysAlaAspAlaLeuValAlaLysA 463
Db 1307 GTGCCATGAAGAGAGAGATATGAAGAAATTCAAAGTTAGGATAAATGGCCTTGTGCTAAGG 1366
Qy 463 laMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrProTrpProGlyAsnA 483
Db 1367 CACAGAAAGTCCCTGAGAGAGGATGGATCATGCAAGATGGCACCATGGCCAGCAAGAAACA 1426
Qy 483 snThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspT 503
Db 1427 ATACTAGGACCATCTCGGAATGATTCAGGTTTTCCTTGGTTCACAGTGGTGGCCTTGATA 1486
Qy 503 hrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheG 523
Db 1487 CCGAGGGTAAATGAGCTTCCCGCCCTGGTCTACGTATCTCGTGAGAAGCGTCTCGGTTC 1546
Qy 523 lnHisHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrA 543
Db 1547 AGCACCAAGAGAGAGCTGGTCCCATGAATGCTGTGTGCTGCTCAGCTGTGCTTACCA 1606
Qy 543 snAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleA 563
Db 1607 ATGGACAGTACATGTTGATCTTGATCTGATCTACATCAACACACAGCAAGGCTCTGC 1666
Qy 563 rglGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysValCysTyrValGlnP 583
Db 1667 GGGAGCTATGCTGCTTCTTATGGACCCCAACCTAGGAAGGAGTGTCTGTTATGTTCAGT 1726
Qy 583 heProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaasnArgAsnThrValP 603
Db 1727 TCCCAACAAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1786
Qy 603 hePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrG 623
Db 1787 TTTTCGATATTAACTTGAGGGGTCTTGATGGCATCCAGGACCAGTTTATGTGGGAACGTG 1846
Qy 623 lYcysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProPro---LysGlyProLysA 642
Db 1847 GTTGTGTTATCAACAGAACTGCTCTATATGGTTATGAACCCCAATTAAGCAGAGAAGA 1906
Qy 642 rgProLysMetValThrCysAspCysCysProCysPheGlyArgLysLys----- 658
Db 1907 AGGGAAGTTTCTGTGCA-----TCATTTTGTGGGGGAGGAAGCAAGCAAGT 1957
Qy 659 -----ArgLysHisAlaLysAspGlyLeuPro- 667
Db 1958 CAAAGAAGAGAGTTCCGGACAAAAGAGTCAACACAGCAGTGGCAGTGTGTCGCAG 2017
Qy 668 -----GluGlyThrAlaAspMetGlyValAspSerAspLysG 680
Db 2018 TTTTCAATCTTCAAGATATAGAGGGGTGTTGAAGGTGCTGGATTTCGATGATGAGAAT 2077
Qy 680 luMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValT 700
Db 2078 CACTTCTTATGTCTCAATAGACTTGGGAAGAGATTTGGCCAGTCTGCAGCGTTTGTTG 2137
Qy 700 hrSerThrLeuMetGluGluGlyGlyValProProSerSerProAlaAlaLeuLeuL 720
```

RESULT 7
AY110415
LOCUS

AY110415 3898 bp mRNA linear HTC 17-OCT-2002

```
2138 CTTCCACTCTCATGGAATATGGTGGTGTCTCAATCTGCAACCCCAAGATCTCTTTTGA 2197
Qy 720 ysgluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluL 740
Db 2198 AAGAAGCTATCCATGATGAAGTGTGGCTATGAGGACAAAGCCGAATGGGAACTGAGA 2257
Qy 740 euGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysA 760
Db 2258 TTGGGTGGATCTACGGTTCGGTACAGAGATATTTCTACTGGATTCAAGATGCATCGC 2317
Qy 760 rglYTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProI 780
Db 2318 GAGGCTGGAGATCAATCTACTGATGCCAAGCGCCAGCTTTCAAGGGGTCTGCTCTTA 2377
Qy 780 leasnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIleP 800
Db 2378 TCAATCTTTCCAGATCGCTTTAAACCAAGTGTTCGGTGGGCACCTTGGTCTCTGTTGA 2437
Qy 800 hePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuG 820
Db 2438 TTTTCAGTCGCCATTTGCCATATGTTAGTACGGCTAT---GGAGGACGCTTAAAGTTC 2494
Qy 820 luArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaT 840
Db 2495 AGAGATTTCGCTACATCAACACCACTATTTATCCATTGACATCGATCCCGCTTCTCTCA 2554
Qy 840 yrcysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerT 860
Db 2555 ACTGTGTTTGGCTGCTATCTGTTGCTCACTGGAAAGTTTCATCTCCAGAGATTAGCA 2614
Qy 860 hrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuG 880
Db 2615 ACTTGTGCTAGTATTTGGTTCATCTCTCTTCATTTCAATTTTGGCCACTGGTATCCTTG 2674
Qy 880 luMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGlnPheTrpValI 900
Db 2675 AGATGAGTGGAGTGGTGTGGCATCGATGAGTGGTGGAGAAATGAACAGTTCTGGGTTA 2734
Qy 900 leGlyGlyValSerAlaHisLeuPheAlaValGlnGlyLeuLysValLeuAlaG 920
Db 2735 TTGAGGTATATCTGCGCATCTTTTTCGGTCTTCCAGGTCTCTCTCAGGTGTTGCTG 2794
Qy 920 lYIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPheA 940
Db 2795 GTATCCGACACCAATTTCACTGTCACTCAAAGGCTCT---GATGAAGATGGTCACTTG 2851
Qy 940 laGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeuLeuIleI 960
Db 2852 CTGAGCTCTACATGTTCAAGTGGACAAACGCTTCTTCATCCCAACCGACCATCTGATCA 2911
Qy 960 leAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerT 980
Db 2912 TTAACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2971
Qy 980 rpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrProP 1000
Db 2972 GGGGACCGCTCTTTGGGCAAGCTCTTCTTGGCTCTGGTGTGTTGTTGTTGTTGTTGTTGTT 3031
Qy 1000 heLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIleTrpSerI 1020
Db 3032 TCCTCAAGGGTCTTATGGGTGGCAAAACCGCATCCGACCATCGATGTTGTTGTTGGGCAA 3091
Qy 1020 leLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgT 1040
Db 3092 TCCCTTCTGGCTTCGATCTTCTCATTTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3151
Qy 1040 hrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3152 TCACCGGCGCCAGATACCCCAACATGTGGTCACTCAACTGC 3189
```

```

DEFINITION   Zea mays CL1166_1 mRNA sequence.
ACCESSION    AY110415
VERSION      AY110415.1 GI:21214824
KEYWORDS     HTC.
SOURCE       Zea mays
ORGANISM     Zea mays

REFERENCE    1 (bases 1 to 3898)
AUTHORS      Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whittitt, M.S.,
              Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
              Maize Mapping Project/DuPont Consensus Sequences for Design of
              Overgo Probes
              Unpublished (2002)
              Coe, E.H.
              Direct Submission
              Submitted (25-APR-2002) Maize Mapping Project, University of
              Missouri, Columbia, MO 65211, USA
COMMENT      If you are interested in getting corresponding physical clones,
              these are publicly available from ZmDB and may be found by BLAST
              searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
              www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
              maize cDNA sequences is either Virginia Walbot, Stanford or Pat
              Schnable, Iowa State, then clones may be requested from ZmDB:
              www.zmdb.iastate.edu.
FEATURES     Location/Qualifiers
              1..3898
                /organism="Zea mays"
                /mol_type="mRNA"
                /db_xref="MaizeDB:630057"
                /db_xref="taxon:4577"
                /clone_lib="Maize Mapping Project/DuPont Cornsensus
                Library"
                /note="this sequence is part of a project of EST
                assemblies resulting from the application of public
                contigs to seed Dupont contigs; this resource was
                assembled by Dupont as part of a collaboration for the
                overgo addressing of BACs in conjunction with the Maize
                Mapping Project"
ORIGIN
Alignment Scores:
Pred. No.:      0          Length:      3898
Score:          3821.50    Matches:      727
Percent Similarity: 75.42%  Conservative: 120
Best Local Similarity: 64.74%  Mismatches: 157
Query Match:     67.66%      Indels:    119
DB:              3          Gaps:      17

US-10-627-132-30 (1-1052) x AY110415 (1-3898)

QY   1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu120
DB   189 ATGGCGGCCAACAGCGGGATGGTGGCGCTCGCAACACCGCAACGAGTTCGTCTATC 248
QY   21 ArgGlyHisGluAsp-----ProLysProLeuArgAlaLeuSerGlyGlnVal 36
DB   249 CGCCACGCGGCGACGCGCGCGCGCGCTAAGCCACGAGAGTGGCAATGGCGAGTTC 308
QY   37 CysGluileCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCys 56
DB   309 TGGCAGATTGGCGGCGACTCTGTGGCGGTTTCAGCCACTGGTGATGTCTTTGTCCTGC 368
QY   57 AsnGluCysGlyPheProValCysArgProCysTyrGluTyrGluArgGluGlyThr 76
DB   369 AATGAGTGGCTTCCCTGTCTGCGCGCTTGTCTATGATGATGATGATGATGATGATGAT 428
QY   77 GlnAsnCysProGlnCysAlaThrArgTyrLysArgLeuLysGlySerProArgValAla 96
DB   429 CAATGCTGCCCTTCAGTGCAGACTAGATACAGACAGACAGAAAGGTAGCCCTCGAGTTCAT 488

QY   97 GlyAspAspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGlu 116
DB   489 GGTGATGAGGATGAGAAAGATGTTGATGACCTAGACAATGAATTCACACTAC----- 539
QY   117 AsnGlnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeu 136
DB   539 ----- 539
QY   137 HisGlyArgMetSerTyrGlyArgGlyPro-----AspAspGlyAsp 150
DB   540 -----AAGCAAGGCGAGTGGGAAGGCCCCAGAGTGGCAGCTGCAAGGAGATGACGCTGAT 593
QY   151 GlyAsnAsnThrPro-----GlnIleProProIleIleThrGlySer 164
DB   594 CTGTCTTCATCTGCTCGCCATGACCCACACCATCGGATTCACAGCTGACAGCGGTCAA 653
QY   165 ArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrGlyHisGly----- 182
DB   654 CAG-----ATATCTGGAGAGATCCCTGATGATCCCTGACCGCTCATTTCTATCCGC 704
QY   183 GluValSerSerSerLeuHisLysArgIleHisProTyrProValSerGluProGlySer 202
DB   705 AGTCCACATCGAGCTATGTTGATCCAGCGTCCCGAGTTCCTGTGAGGATTTGTG----- 758
QY   203 AlalysTrpAspGluLysGlu-----ValSerTrpLys 214
DB   759 -----GACCCCTCGAGGACTTGAATTCCTATGGGCTTAATAGTGTGACTGGAAG 809
QY   215 GluArgMetAspAspTrpLysSerLysGln----- 224
DB   810 GAAAGAGTTGAGAGCTGGAGGGTTAAACAGGACNNNNNTATGTTGCAAGTGACTAATAA 869
QY   225 -----GlyIleLeuGlyGlyAlaAspProGluAspMetAsp 237
DB   870 TATCCAGAGGCTAGAGGAGGACATGAGGGGAGTGGCTCAATGGAGAGATATGCAA 929
QY   238 AlaAspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAla 257
DB   930 -----ATGTTGATGATGACGCGCTACCTTTGAGCGCGCATCGTGCATATTC 977
QY   258 SerSerLysValAsnProTyrArgMetValIleValValArgLeuValAlaPhe 277
DB   978 TCAACACGAGCTCAACCTTTACCGGTAGTATCATCTCCGCTTATCATCTCGCTTC 1037
QY   278 PheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyIleTrpLeuValSer 297
DB   1038 TTCTTCCAGTATCGTATCAGTCCATCGAGTGGTGTATGATGATGATGATGATGATGAT 1097
QY   298 IleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPhe 317
DB   1098 GTTATCTGCGAGGTCTGGTTTGGCTTCTGCTGGCTTCTAGATCAGTCCCAAAATGGTAT 1157
QY   318 ProIleAspArgGluThrTyrLeuAspArgLeuSerLeuLeuArgTyrGluArgGluGlyGlu 337
DB   1158 CCAATCAACCGTGAGACATATCTGACAGGCTTCATTTGAGGTATGATAGAGAGGAGAG 1217
QY   338 ProSerLeuLeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluPro 357
DB   1218 CCATCAGAGCTGGCTCCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1277
QY   358 ProLeuValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLys 377
DB   1278 CCATGATCAGAGCAACACACTGTTTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1337
QY   378 ValSerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGlu 397
DB   1338 GTGTCATGCTATGTTCTGATGATGGCTCAGCTATGCTGCTGCTGCTGCTGCTGCTGCTG 1397
QY   398 ThrAlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArg 417
DB   1398 ACCGCCGAAATTGCTAGAAAGTGGTTCCTTTTGTAGAGAGCAATATTTGAAACCAAGA 1457
QY   418 AlaProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThr 437

```

```
Db 1458 GCTCAGAAATTTACTTTGCTCNNNNNNTAGATTACCTGAAGGACAAAATTCACACCTTCA 1517
Qy 438 PheValGlnGluArgAlaMetIysArgGluTyrGlnGluPheIysValArgIleAsn 457
Db 1518 TTTGTGAAGGAACGCGCAATGAAGAGGAGATGAAGAAATTCAAAGTAAGAAATCAAT 1577
Qy 458 AlaLeuValAlaLysAlaMetLysValProAlaGluGlyTyrIleMetLysAspGlyThr 477
Db 1578 GCCCTTGTTCGAAGACACAGAAAGTGCCTCAAGAGGGGTGGACATGGCTGATGGAACT 1637
Qy 478 ProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHis 497
Db 1638 GCTTGGCTGGCAATAACCTTAGGCACCATCTCGCATGATTCAGGTGTTCTTGGGGCAC 1697
Qy 498 SerGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlu 517
Db 1698 AGTGGTGGCTCGACACTGAGGAAATGAGTTACACAGCTTGTGTCTATGTCTCTCGTGAA 1757
Qy 518 LysArgProGlyPheGlnHisIleLysLysAlaGlyAlaMetAsnAlaLeuIleArgVal 537
Db 1758 AAGAGACAGAGCTTTCAGCATCACAGNAGGCTGGTGCATGAATGCCTGATTCGTGTA 1817
Qy 538 SerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsn 557
Db 1818 TCTGCTGTGTGACAAATGGTGCCTATCTCTCAATGGGATTCGACCATTACTTCAAT 1877
Qy 558 AsnSerIysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLys 577
Db 1878 AGCAGCAAGAGCTCTAGAGAGCAATGTGCTTCATGATGGATCCGGCTCTAGGAAGGAAA 1937
Qy 578 ValCysTyrValGlnPheProGlnPheArgPheAspGlyIleAspValHisAspArgTyrAla 597
Db 1938 ACTTGTATGTACAAATTTCCACAGAGATTTGATGGCATTGATTCGCAGCATCGATATGCT 1997
Qy 598 AsnArgAsnThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyPro 617
Db 1998 AATCGGAACATAGTCTCTTTGATATCAACATGAAGGCTCTGGATGGCATTTCAGGGTCCA 2057
Qy 618 ValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnPro--- 636
Db 2058 GTCTACGTGGGAACAGGATGCTGTTTCAATAGGCAGGCTTTGTATGGATACGATCCCTGTT 2117
Qy 637 ---ProLysGlyProLysArgProLysMetValThrCysAspCys----- 650
Db 2118 TTGACTGAAGCTGATCTGGAGCTTAACATTGTTGTAAAGAGCTGCNNNNNNNNNNNN 2177
Qy 650 ----- 650
Db 2178 NNNNNNNNNNNNNNATGATAGTCAAAAGCCGTATTATCAAGAGAACAGAAATCTTCA 2237
Qy 651 CysProCysPheGlyArgLysLysArgLysHisAlaLysAspGlyLeuProGluGlyThr 670
Db 2238 GCTCCCATCTTCAAC-----ATGGAAGACATCGAGGAGGGTATT 2276
Qy 671 AlaAspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLys 690
Db 2277 GAA-----GGTTACAGAGTGAAGAGGTGATGCTTATGTTCCAGAGGAAATTCGAGAAA 2330
Qy 691 ArgPheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGluGlyValPro 710
Db 2331 CGCTTGTGCTAGTCTCCAATCTTCATTCGATCCACCTTTATGACTCAAGGTGGCATACCA 2390
Qy 711 ProSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 730
Db 2391 CCTTCAACAAACCCAGCTCTCTACTAAAGGAAGCTATCCATGTCATCAGTTGTGGATAT 2450
Qy 731 GluAspLysThrAspTrpGlyLeuGluGlyTyrIleTyrGlySerIleThrGluAsp 750
Db 2451 GAGGACAAACTGAATGGGAAAAGAGATGGCTGGATCTATGTTTCAGTTACGGAGGAT 2510
Qy 751 IleLeuThrGlyPheLysMetHisCysArgGlyTyrArgSerValTyrCysMetProLys 770
```

```
Db 2511 ATTCTGACTGGGTTTAAANNNNNNNNAGGGCTGGCAATCAATCTACTCATCGATCCACCA 2570
Qy 771 ArgAlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 790
Db 2571 CGACCTTGTTCGAAGGGTCTCGACCAATCAATCTTCTGATCGTCTTATCAGGTGCTC 2630
Qy 791 ArgTrpAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGly 810
Db 2631 CGTTGGGCTCTTGGGTCAGTGGAAATTCCTCTTAGCAGACATTTGCTCTATCTGATGTT 2690
Qy 811 TyrIysAsnGlyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrIleTyr 830
Db 2691 TAC----AATGGACGATTTGAAGCTTTTGGAGAGGCTGGCTTACATCAACACCATTTAT 2747
Qy 831 ProPheThrSerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLysLeuThr 850
Db 2748 CCAATCACATCTGTTCCGCTTATCGCCTATTTGTGTCTCTGCTATCTCTCTTACC 2807
Qy 851 GlyLysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPhe 870
Db 2808 AATAAATTTATCATCTCTGAGATTAGCAATATATGCTGGGATGTTCTTCAATCTCTTTC 2867
Qy 871 MetSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluGlu 890
Db 2868 GCCTCCANNNGCCACTGGTATATTGGAGCTTAGATGGAGTGGTGTGGCATTGAAGAT 2927
Qy 891 TrpTrpArgAsnGluGlnPheTrpValIleGlyValSerAlaHisLeuPheAlaVal 910
Db 2928 TGTGGGAAATGAGCAGTTTTGGTTATTGGTGGACCTCTGCCCATCTCTTCGCGATG 2987
Qy 911 ValGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLys 930
Db 2988 TTCAGGGTCTCTGTGAAGTGTGGCTGGGATTTGATACCAACTTCACAGTTACCTCAAAG 3047
Qy 931 AlaThrGlyAspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrLeu 950
Db 3048 GCATCT---GATGAGGATGGCAGCTTTGCTGAGCTATATGTTTCAAGTGGACAGTTG 3104
Qy 951 LeuIleProThrThrLeuLeuIleAsnValIleGlyValValAlaGlyIleSer 970
Db 3105 CTCAATCTCGACCATCTTCTTGTCTATTAACTGCTGGATGGTGGCAGGAAATTCG 3164
Qy 971 AspAlaIleAsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAla 990
Db 3165 TATGCCATTAACAGTGGCTACCAATCTCGGCTCCGCTCTTTGGAAAGCTTCTTCTCG 3224
Qy 991 PheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArg 1010
Db 3225 ATCTGGGTGNTCTCCCATCTGTACACGTTCTCAAGGGTCTCATGGGAAGGACAAACGC 3284
Qy 1011 ThrProThrValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuTrp 1030
Db 3285 ACACCAACATCTGTCATGTCGTGCTCATCTTCTTGCATCTATCTTCTCTCTGCTGG 3344
Qy 1031 ValArgIleAspPropheIleValArgThr---LysGlyProAspValArgGlnCysGly 1049
Db 3345 GTGAAGATCGATCTTTTCATCTCCCGCACACAGAAAGCTGTGCTTGGGGCAATGTGGC 3404
Qy 1050 IleAsnCys 1052
Db 3405 GTCAACTGC 3413
RESULT 8
CNS0A1ML 3911 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSUTPGH52A06 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION BX832166
VERSION BX832166.1 GI:42459101
KEYWORDS HTC; SUT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
AUTHORS
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.

TITLE
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

JOURNAL
Unpublished

REFERENCE
AUTHORS
2 (bases 1 to 3911)

TITLE
Direct Submission

JOURNAL
Genoscope.

REFERENCE
AUTHORS
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr)

TITLE
Web : www.genoscope.cns.fr)

COMMENT
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EP/Full_length
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

```

/ gene= AC596474.0
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3911
Score: 3719.00 Matches: 711
Percent Similarity: 76.57% Conservative: 132
Best Local Similarity: 64.58% Mismatches: 193
Query Match: 65.85% Indels: 65
DB: 3 Gaps: 17

US-10-627-132-30 (1-1052) x CNS0A1ML (1-3911)
QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgHcnGluLeuValLeu 20
Db 430 ATGAACACCCGGTGGTTCGGTTAATCGCCGGTCTCACAAACAGGAATGAGTTGGCTCAT 489
QY 21 ArgGlyHisGluAspPro-----LysProLeuArgAlaLeuSerGlyGlnValCysGlu 38
Db 490 AATGCCGATGAGAATGCCGCAATAAGATCAGTCCAGAGCTGAGTGACAGACATGTCAA 549
QY 39 IleCysGlyAspGlnValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGlu 58
Db 550 ATCTGACAGATGAGATCGAATTGACTGTGTGATGGAGAACCGGTTGTGGCATGTAAACGAA 609
QY 59 CysGlyPheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsn 78
Db 610 TGTGCATTCCCTGTGTAGACCTTGTCTATGAGTAGTCGAAAGACGAGNAGCAATCAAGCT 669
QY 79 CysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAsp 98
Db 670 TGTCCACAGTGCAAAACCCGTTTCAAACGCTCTAAAGGAAGTCCAAGAGTTGAAGGTGAT 729

```

99 AspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGln 118
 730 GAAGAGGAAGATGACATTTGATGATTTAGACAATGAGTTTGAGTATGCAAAAT 780
 119 GlnArgGlnLeuGluGlyAsnMetGlnAsnSerClnIleThrGluAlaMetLeuHisGly 138
 781 -----AATGGATTTGGATTGGATCAGGTTTCTGAAGTATGTCAATCTCT 825
 139 ArgMetSerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrPro----- 155
 826 CGTCGCAACTCCGTTTCCACCAATCGATTGGAT-----TCAGCTCCACCTGGCTCT 879
 156 GlnIleProProIleThrGlySerArgSerValProValSerGlyGlu----- 172
 880 CAGATTTCCATTGCTGATTACCGGCGACGAGGAGCTTGAGATTTCTCTGATAGACATGCT 939
 173 -----PheProIleThrAsnGlyTyrGlyHisGly---GluValSerSer-SerLeuHi 189
 940 CTTATTGTTCTCTCTCATTCTGTT---GGTCATGGCATAGAGTTATCTCTGTTCTCTTTC 996
 189 sLys-----ArgIleHisProTyrProValSerGluProGlySerAlaIysTrpAs 206
 997 TGACCCGACCGTGGCTGCACATCCAGGCTATGGTACCTCAGAAAGATCTTGGGTTTA 1056
 206 pGluLysLysGluValSerTrpLysGluArgMetAspAspTrpLysSerLysGlnGly-- 225
 1057 TGGTTATGGAAGTGTGCTTGGAAAGATCGATGGAGGAATGGAAGAGAAGCAGAATGA 1116
 226 -----IleLeuGlyGlyAlaAspPro-----GluAspMetAspAlaAs 239
 1117 GAAACTTCAGGTTGTAGGCGATGAAGAGAGATCCTGATTTTGAAGATGGTGATGCTGA 1176
 239 pValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSerSe 259
 1177 TTTTCCAAATGATCGATGAGGAGGACGACCATTTGCTGGAAGATACCAATCAATCGAG 1236
 259 rLysValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPhePheLe 279
 1237 CAAGATAAATCCCTTACCGGATGTTAATTTGTGTACGTCTTGATTTCTGGTCTCTTCTT 1296
 279 uArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerIleIl 299
 1297 TCATAACCGATTATCTTCAACCCGTCAAAGATGATGATGCTTTGTGGCTTATTTCTGTAT 1356
 299 eCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPheProIl 319
 1357 ATGTGAGATATGTTTGTCTGTTTCATGGGTTCTTGATCAGTTCCCTTAATGGTACCCTAT 1416
 319 eAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGluProSe 339
 1417 CGAGCGAGAAACGATCTTGACCGACTCTCATTAAGATATGAGAAAGAGGAAACCGTC 1476
 339 rLeuLeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluProProLe 359
 1477 GGAACATATCCCTGTCGATGTTATTTGTGTAGTCAGTGGATCCATTCGAAAGAGCTCCGCT 1536
 359 uValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysValSe 379
 1537 TATTACTGCAATATCTGTTCTGTCTTATTTCTGCTGTGTGATTATTCCTGTGATAGTTGC 1596
 379 rCysTyrValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAl 399
 1597 TTGTTACGTTATCTGATGATGGTGTCTATGCTTACTTTCGAAGCTCTTTCTGAGACCGC 1656
 399 aGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArgAlaPr 419
 1657 TGAATTCGCAAGAAATGGGTTCTTTCTGCAAGAAATATTGTATTGAGCCTCGTCTCC 1716
 419 oGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThrPheVa 439
 1717 CGAATCGTATTCTGCCATAAATGGACTACTTGAAGAATAAAGTTTCATCCCGCATTTGT 1776
 439 lGlnGluArgArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAlaLe 459

```
Db 1777 TAGGAGCGCGAGCCATGAAGAGAGATTATGAAGAAATTCAAAAGTAAGATCAATGCTTT 1836
Qy 459 uValAlaLysAlaMetLysValProAlaGluGlyTTPleMetLysAspGlyThrProTr 479
Db 1837 AGTAGCAACAGACAGAAAGTGCCTGAAGATGGGTGGAGCTATGCAAGACGGTACACCTTG 1896
Qy 479 pProGlyAsnAsnThrArgAspHisProGlyMetIleGlnVal-PheLeuGlyHisSerG 499
Db 1897 GCCCGTAATAGTGTGGAGATCATCTCGCATGAATCAGGTCTTTCTTGGAAAGTGACC 1956
Qy 499 lyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysA 519
Db 1957 GTGTTCTGTGTGTCGAAACACACAGGTTCCTCGATTAGTTTACGTTTCTCGTGAGA 2016
Qy 519 rgProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSerA 539
Db 2017 GACCGGATTTGATCACCATAAGAGGCTGGAGCTATGAATTCCTCGATACGAGTCTCTG 2076
Qy 539 laValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnS 559
Db 2077 GGGTTCATCAAAATGCTCTTACCTTCCTGAATGTCGATTGTGATCATTACATCAACATA 2136
Qy 559 erLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysValC 579
Db 2137 GCAAGCTCTTAGAGAAGCAATGTGTTTCATGATGGATCCTCAGTCAGGAAGAAATCT 2196
Qy 579 ySerValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsnA 599
Db 2197 GTTATGTTTCAGTTCCTCAAAAGTTTCGATGGATTTGATGAGCAGATCGATCTCAATC 2256
Qy 599 rgAsnThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProValT 619
Db 2257 GCAATGTTGTGTTCTTTGATATCAATATGAAGGTTTGGATGGGCTACAGGGGCTATAT 2316
Qy 619 yrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProProLysG 639
Db 2317 ACGTCGGTACAGTTGTGTTTCAGGAGCAAGCGCTTTACGGATTTGATGCACCGAAGA 2376
Qy 639 lyProLysArgProLysMetValThrCysAspCysCysPro-----CysP 654
Db 2377 AGAAGAGGGGCCACGTAAG---ACATGCAATTTGTCGCCAAATTTGTTCTCTCTATGTT 2433
Qy 654 heGly---ArgLysLysArgLys----- 660
Db 2434 TTGGTTCAAGAAAGAAATCGTAAGACCAAGACAGTGGCTCGGATGAAGAAAGAAATA 2493
Qy 661 -----HisAlaLysAspGlyLeuProGluGlyThrAlaAspMetG 674
Db 2494 GGGAGCGCTCAAGCAGATCCACGCAATTAGAAATATCGAAGAGGGCGCGTCACTAAAG 2553
Qy 674 lyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyG 694
Db 2554 GTTCTAAGCTAGACAGTCAACCGGCAATGCAATGAAGTTGGAGAAGAAATTTGGGC 2613
Qy 694 lnSerAlaAlaPheValThrSerThrLeuMetGluGlyGlyValProProSerSerS 714
Db 2614 AGTCTCTGTATTGTTGTCATCTGCGGTATGGAGAATGGTGGATGGCTAGAACGCA 2673
Qy 714 erProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysT 734
Db 2674 GCCCGCTGTCTGCTTAAAGAGCAATCCCAAGTCATTAGTTGCGGATATGAAGATAAA 2733
Qy 734 hrAspTrpGlyLeuGluLeuGlyTTPleTyrGlySerIleThrGluAspIleLeuThrG 754
Db 2734 CTGAATGGGAAAGAGATGGGTGGATCTATGTTCTCTGTTACCGAAGATATTTCTACGG 2793
Qy 754 lyPheLysMetHisCysArgGlyTTPArgSerValTyrCysMetProLysArgAlaAlap 774
Db 2794 GTTTTAAGATGCTCTCTATGTTGGAGATCTGTTTATTTGATACACCAAGATTAGCGGCTT 2853
Qy 774 heLysGlySerAlaProIleAsnLeuSerArgLeuAsnGlnValLeuArgTTPAlaL 794
```

```
Db 2854 TCAAAGGATCAGCTCCAATCAATCTTTTCGATCGTCTCCATCAAGTTCTTCGATGGCGC 2913
Qy 794 euGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnG 814
Db 2914 TTGGGTCGGTTGAGATTTTCTTGAGTAGGCAATGCTCTATTTGGTATGGTTAT---GGAG 2970
Qy 814 lyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrIleTyrProPheThrS 834
Db 2971 GTGGGTTGAATGGCTTGAGCGGTGTCCTTACATTAACCTCTGTGGTTTACCCGTTGACCT 3030
Qy 834 erLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheI 854
Db 3031 CTCTACCGCTCATCGTTTACTGTTCTCTCCCTGCCATCTGCTCTTCACCTGGAATAATCA 3090
Qy 854 leMetProSerIleSerThrPheAlaSerLeuPheIleAlaLeuPheMetSerIleP 874
Db 3091 TCGTTCGCCGAGATTAGCACTATGCGAGTATCTCTTCATGGCGCTCTTCTGTCGATTG 3150
Qy 874 heAlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgA 894
Db 3151 CAATAACGGGTATTTCTCGAGATGCAATGGGCAAGTTGGGATCGATGTTGGTGGAGAA 3210
Qy 894 snGluGlnPheTrp-VallleGlyGlyValSerAlaHisLeuPheAlaValGlnGly 913
Db 3211 ACGAACAGATTTTGGGGTTCATTGGAGGTGTTTCTCGCATCTGTTGCTCTCTTCCAAGGT 3270
Qy 914 LeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGly 933
Db 3271 CTCTCAAGGTCTTGTGTTGTCGACACTTAACCTTCACATGCATCACTCAAAAGCAGCT--- 3327
Qy 934 AspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIlePro 953
Db 3328 ---GATGATGGAGAGTTCTCTGACCTTTTACCTTCAAATGGACTTCACTTCTCATCCCT 3384
Qy 954 ProThrThrLeuLeuIleIleAsnValIleGlyValAlaGlyIleSerAspAlaIle 973
Db 3385 CCAATGACTCTACTCATATAAAGCTCATTTGGAGTCATAGTCGGAGTCTCTGATGCCATC 3444
Qy 974 AsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPheAlaPheTrpVal 993
Db 3445 AGCAATGGATGACACTCGTGGGACCGCTTTTTCGGAAGACTGTTCTTTGCACTTGGGTC 3504
Qy 994 IleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThr 1013
Db 3505 ATCATTCATCTTTACCGCTTCTTAAAGTTTGTCTGGGAAACAAAGATAGAATGCCAAC 3564
Qy 1014 ValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIle 1033
Db 3565 ATTATTGCTGCTGCTCCATCCCGCTCGACTCTTACACTTCTTTGGGCCCGGGTT 3624
Qy 1034 AspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3625 AATCCGTTTGTGGCGAAA---GGCGGTCCTATTCTCGAGATCTGTGTGTAGACTGC 3678
RESULT 9
AV112236 3728 bp mRNA linear HTC 17-OCT-2002
LOCUS Zea mays CL1160_1 mRNA sequence.
DEFINITION AY112236
VERSION AY112236.1 GI:21216826
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3728)
AUTHORS Hainey, C. F., Dolan, M., Miao, G. H., Vogel, J. M., Whitsitt, M. S.,
Arthur, L. W., Hanafey, M., Morgante, M. and Tingey, S. V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 3728)
```

AUTHORS TITLE JOURNAL

Coe, B.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES

Location/Qualifiers

1..3728
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:630049"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3728
Score: 3633.50 Matches: 703
Percent Similarity: 75.38% Conservatives: 99
Best Local Similarity: 66.07% Mismatches: 211
Query Match: 64.33% Indels: 51
DB: 3 Gaps: 12

US-10-627-132-30 (1-1052) x AV112236 (1-3728)

QY 37 CysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCys 56
DB 292 TCCAGATCGCGGATCGCGTGGGACACTAGCGCGGAGGAGACGTCTTACCGCGTGC 351
QY 57 AsnGluCysGlyPheProValCysArgProCysTyrGluTyrGluArgGluGlyThr 76
DB 352 GAGCTGTGGGGTTCGGGTGCGCCCTCTAGAGTACGAGGAGGAGGAGGAGGAGGAG 411
QY 77 GlnAsnCysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAla 96
DB 412 CAAGCGTGNNNNAGTGNNNNCAAGTACAGCCCAAGCGGAGTCCAGCGATCGA 471
QY 97 GlyAspAsp---AspGluAspIleAspAspLeuGluHisGluPheAsnIle----- 113
DB 472 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 528
QY 114 ---AspAspGluAsnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThr 132
DB 529 TCTGCAATGACACAGACAGAGAGATTGCTGACAGGATCGGAGTGGCGGATGAAT 588
QY 133 GluAlaMetLeuHis-----GlyArgMetSerTyrGlyArgGlyProAspAspGlyAsp 150
DB 589 GCTNNNNCAGCGGGGATGTTGGCGCCCAAGTATGACAGTGGTACAGTCCGGCTACC 648
QY 151 GlyAsnAsnThrProGlnIlePro---ProIleIleThrGlySerArgSerValProVal 169
DB 649 AAGTACGACAGTGGTGGAGATCTCTCGGGGATACATCCNNNNNNNNNNNNNNNN 708
QY 170 SerGlyGluPheProIleThrAsnGlyTyrGlyHisGlyGluValSerSerSerLeuHis 189
DB 709 TCGGAGAAATCCCTGGTCTCCCTGACCATATATGATGCTCTCTACTGGGAACATT 768
QY 190 LysArgIleHisProTyrProVal-----SerGluProGlySerAlaLysIle 205
DB 769 GCGAGCGCGCCCATTTCCCTATATGAATCATTCATCCGTCGAGGGAATCTCT 828

QY 206 AspGluLysLysGluValSerTrpLysGluArgMetAspAspTrpLysSerLysGln---- 224
DB 829 GGTAGCGTGGGAATGTTTGCCTGGAAGAGAGAGGTTGATGGCTGGAATAAGACGAGAC 888
QY 225 ---GlyIleLeu-----GlyGlyGlyAlaAspPro 233
DB 889 AAGGAGACAAATTCCTCATGACGAATGGCACACGATTCCTCTGAGGGCGCGGGTGT 948
QY 234 GluAspMetAspAla-----AspValAlaLeuAsnAspGluAla 246
DB 949 GGTGATATTGATGATCACTCAATGATGGAAGATGCTTATTAACCATGAACT 1008
QY 247 ArgGlnProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMet 266
DB 1009 CGCCAGCCTCTATCTAGGAAGTTCCACTTCTCTCCAGGATAAATCATACAGATG 1068
QY 267 ValIleValValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisPro 286
DB 1069 GTCATGTGTACGATGATGTTCTAAGCATCTTCTTGCACTACCGGATCAAAATCCT 1128
QY 287 ValProAspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIle 306
DB 1129 GTGCGTAATGCATACCCACTGTGGCTTCTATCTGTATATGTAGATCTGGTTGTCT 1188
QY 307 SerTrpIleLeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAsp 326
DB 1189 TCCTGGATATTGGATCAGTTTCCAAAGTGGTTTCCAATCAACCGGAGACTTACCTGAT 1248
QY 327 ArgLeuSerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeu 346
DB 1249 AGACTCGCATTAAGGTATGACCGGAGAGGTAGCCATCTCAGTTGGCTGTCTTGACANN 1308
QY 347 PheValSerThrValAspProLeuLysGluProLeuValThrAlaAsnThrValLeu 366
DB 1309 NNNGTCACTACTGTCGACCAATGAAGGAGCTCTCTTGTACCTGCCAATACCTGTCTA 1368
QY 367 SerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspGly 386
DB 1369 TCCATTCCTCGTGGACTATCTCTGTGTAAGGTCTCTTGTATGTATCTGATGATGA 1428
QY 387 AlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpVal 406
DB 1429 GCTGTATGTGACATTTGATGCTAGCTAGACTTACAGACTTTCAGAGTTTGTAGAAATGG 1488
QY 407 ProPheCysLysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLys 426
DB 1489 CCATTTTGAAGAGTACAACTAGAGCTCTGTAAGTGTCTCTCTCCAGAA 1548
QY 427 ValAspTyrLeuLysAspLysValGlnProThrPheValGlnGluArgAlaMetLys 446
DB 1549 ATTGATTACTTGAAGGACAAAGTGCACCTTCAATTTGTTAAAGACCGCGGNNNNNN 1608
QY 447 ArgGluTyrGluGluPheLysValArgIleAlaLeuValAlaLysAlaMetLysVal 466
DB 1609 NNN 1668
QY 467 ProAlaGluGlyTyrIleMetLysAspGlyThrProTyrProGlyValAsnThrArgAsp 486
DB 1669 CCTGAGGAAGGATGGATCATCAAGATGGCACCATGCGCCAGGAAACAATACAGGAC 1728
QY 487 HisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsn 506
DB 1729 CATCTGGAATGATTCAGGTTTCTCTGGTGCACAGTGGTGTCTTGATCTAGAGGAT 1788
QY 507 GluLeuProArgLeuValTyrValSerArgLysArgProGlyPheGluHisLys 526
DB 1789 GAGTACCCCTTGTGTTCTCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1848
QY 527 LysAlaGlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaPro 546
DB 1849 AAGCTGTGGCATGAATGCTCTTGTGCGCTCTCAGCTGTCTTACCAATGACAAATAC 1908
QY 547 MetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMet 566


```
Db 1909 ATGTTGAATCTTGATTTGATCACTACATCAACACAGTAAGGCTCTCAGAGGAGCTATG 1968
Qy 567 CysPheLeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArg 586
Db 1969 TGCTTCCTTATGGATCTCAACCTAGGAGGAGTGTCTGCTATGTTCAAGTTTCCCCAGAG 2028
Qy 587 PheAspGlyLeuAspValHisAspArgTyrAlaAsnArgAsnThrValPheAspIle 606
Db 2029 TTTGATGGTATTGATAGGAATGATCGATATGCCAACAGGAACACCGTGNNNNCGATATT 2088
Qy 607 AsnMetLysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPhe 626
Db 2089 AACTTGAGAGCTCTTGATGGCATCCAAGGACCACTTTATGTGGGCACTGGCTGTGTTTC 2148
Qy 627 ArgArgGlnAlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetVal 646
Db 2149 AACAGAACAGCTTATATGTTTATGAGNNNNNAATTAAGCAAAAGAGGGTGGTTTCTTG 2208
Qy 647 ThrCysAspCysCysProCysPheGlyArgLysLysArgLysHisAlaLysAsp- 664
Db 2209 TCATCACTATGTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2268
Qy 665 -----GlyLeuProGlu 668
Db 2269 NNGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTAGAGGAG 2328
Qy 669 GlyThrAlaAspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPhe 688
Db 2329 GGAGTTGAAGCGCTGGGATTTGATGATGAGAAATCACTCTTATGTTCTCAAAATGAGCTG 2388
Qy 689 GluLysArgPheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGluGly 708
Db 2389 GAGAAGAGATTGGCCAAATCTGCAGCTTTTGTGCGTCCACTGATGGAATNNNNNNNN 2448
Qy 709 ValProProSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCys 728
Db 2449 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTATCAAGTTGT 2508
Qy 729 GlyTyrGluAspLysThrAspTyrGlyLeuGluLeuGlyTyrIleTyrGlySerIleThr 748
Db 2509 GGCTACGAGGACCAAGATTGAATGGGGAAGTGGGATGGGATCTATGTTCTGTGACG 2568
Qy 749 GluAspIleLeuThrGlyPheLysMetHisCysArgGlyTyrArgSerValTyrCysMet 768
Db 2569 GAAGATATTCATCTGGGTTCAAGATGCACGACGAGGCTGGCGGTGCATCTACTGCAATG 2628
Qy 769 ProLysArgAlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGln 788
Db 2629 CCTAAGCGCGCGCTTCAAGGGATCGGCTCCCATCAATCTCTCAGACCGTCTGAACCA 2688
Qy 789 ValLeuArgTyrAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeu 808
Db 2689 GTGCTCCGGTGGGCTCTCGGTTCACTGGTGAATATCTCTTTTCAGCGGCAATGNNNTATGG 2748
Qy 809 TyrGlyTyrLysAsnGlyAsnLeuLysTyrLeuGluArgPheAlaTyrIleAsnThrThr 828
Db 2749 TACGGGTAC---GGAGGAGCGCTGAAGTCTTGGAGAGATTCGCCCTACATCAACACCA 2805
Qy 829 IleTyrProPheThrSerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeu 848
Db 2806 ATCTACCGCTCAGCTCCCTCCGCTCCTCATTTACTGTATCTCTGCCCTGCCATCTGCCTG 2865
Qy 849 LeuThrGlyLysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIleAla 868
Db 2866 CTCACGGGGAAGTTATCATCCAGAGATCAGCAACTTCGCTAGTATCTATGTTCTATCTCT 2925
Qy 869 LeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValSerIle 888
Db 2926 CTCCTCATCTCCATCTTCCCAAGGATATCTTGGAGATGAGTGGAGCGGCTGGGCATC 2985
Qy 889 GluGluTyrTyrArgAsnGluGlnPheThrValIleGlyGlyValSerAlaHisLeuPhe 908
:::|
```

```
2986 GACGAGTGGTGGAGGAACGACGACGAGTTCTGGGTTCATCGGAGGCATCTCCGCCACCTCTTC 3045
Qy 909 AlaValValGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThr 928
Db 3046 GCCGTCTTCCAGGCGCTCTCTCAAGGTGCTTCCGGGATCGACACCAACTTCACCGTCACC 3105
Qy 929 SerLysAlaThrGlyAspGluAspGluPheAlaGluLeuTyrAlaPheLysTyrThr 948
Db 3106 TCCAAGGCGCTCG--GATGAAGACGCGGACTTCGCGGAGCTGTACATGTTCAAGTGGACG 3162
Qy 949 ThrLeuLeuIleProProThrThrLeuLeuIleAsnValIleGlyValValAlaGly 968
Db 3163 ACATCTTGTATCCCGCCACCACTCTCATCAACCTGTGTGGCGTGTGTGTCGCGG 3222
Qy 969 IleSerAspAlaIleAsnGlyTyrGlnSerTyrGlnProLeuPheGlyLysLeuPhe 988
Db 3223 ATCTCTACGCCATCAACAGCGGTACCACTGTGGGTCCGCTCTTCGGCAAGCTCTTC 3282
Qy 989 PheAlaPheThrValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGln 1008
Db 3283 TTCGCTTCTGGGTGATCGTTCACTTGTACCGGTTCCTCAAGGCTCTCATGGGTGCGCAG 3342
Qy 1009 AsnArgThrProThrValValIleThrSerIleLeuLeuAlaSerIlePheSerLeu 1028
Db 3343 AACCGCACCCCGACCACTCGTGTGTTGCTGGGCGATCTCTGGCGTGCATCTTCTCTCTG 3402
Qy 1029 LeuTyrValArgIleAspProPheIleValArgThrLysGlyProAspValArgGlnCys 1048
Db 3403 CTGTGGGTTCGCATCGATTCGTTCAACCAACGGGTCACTGCGCGGATATCTCGAACCTGT 3462
Qy 1049 GlyIleAsnCys 1052
Db 3463 GGCATCAACTGC 3474

RESULT 10
LOCUS AY110079 3696 bp mRNA linear HTC 17-OCT-2002
DEFINITION Zea mays CL1164_1 mRNA sequence.
ACCESSION AY110079
VERSION AY110079.1 GI:21214162
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3696)
AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 3696)
AUTHORS Coe, E.H.
JOURNAL Direct Submission
TITLE Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
LOCATION/Qualifiers
1..3696
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:630055"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST

FEATURES
source
```



```
Db 2165 GCCTATATGGTATGAGCTCCAGTCAAGNNNNNNNGCCAGGCTTCTTCTCTCGCTT 2224
Qy 650 CysCys-----ProCysPheGlyArgLysLysArg 659
Db 2225 TGTNNNNNAAGNNNNNGACGTCAAATCTAAGAAGAGCTCGGAAAAGAAAGTCCACAT 2284
Qy 660 LysHisAlaLysAspGlyLeuPro-----GluGlyThrAla 671
Db 2285 AGACACGACAGACAGCTTCTGTACCAAGTATTTAATCTCGAAGATATPAGAGGAGGATGAA 2344
Qy 672 AspMetGlyValAspSerAspLysGluMetSerHisMetAsnPheGluLysArg 691
Db 2345 GGTTCTCAGTTTGATGATGAGAAATCGTGATTAATGCTCAATGAGCTTGGAGAGAGA 2404
Qy 692 PheGlyGlnSerAlaPheValThrSerThrLeuMetGluGluGlyValProPro 711
Db 2405 TTTGGCCAGTCCAGTGNNNNGTAGCCTCTACTCTGATGGAATATGTTGGTGTTCACAA 2464
Qy 712 SerSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyThrGlu 731
Db 2465 TCTGCAACTCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2524
Qy 732 AspLysThrAspTTPGlyLeuGluGlyTTPilleTyrGlySerIleThrGluAspIle 751
Db 2525 GACAAACTGACTGGGNACTGAGATTGGGTGGATCTATGGTCTGTGTACAGAGACATT 2584
Qy 752 LeuThrGlyPheLysMetHisCysArgGlyTTPArgSerValTyrCysMetProLysArg 771
Db 2585 CTCACCGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2644
Qy 772 AlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArg 791
Db 2645 CCAGCTTTCAAGGAGATCTGCTTATCAACTTTCGGATCGTGTGAATCAAGTGCCTCGG 2704
Qy 792 TTPAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyr 811
Db 2705 TGGGCTCTTGGTTCATTGAATCTTTTCAGCAGGCATTTGCCCATATGATGATGGCTAT 2764
Qy 812 LysAsnGlyAsnLeuLysTTPLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 831
Db 2765 --GGAGCCCGCTTAAATCTCTGGAGAGATTGCTTATATCAACACACAAATTTATCCA 2821
Qy 832 PheThrSerLeuProLeuAlaTyrCysThrIleProAlaValCysLeuLeuThrGly 851
Db 2822 CTCACATCAATCCCGCTCTCTGCTGCTGCTATATGCGCAGCAGTTGTCTCTCACTGG 2881
Qy 852 LysPheIleMetProSerIleSerThrPheAlaSerLeuPheIleAlaLeuPheMet 871
Db 2882 AAGTTCATCATCCCAAGATTAGTAACCTAGAGAGTGTGGTTTATATCGCTCTTTATC 2941
Qy 872 SerIlePheAlaThrGlyIleLeuGluMetArgTTPSerGlyValSerIleGluIleTTP 891
Db 2942 TCAATCTTTGCCACTGGTATCTCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3001
Qy 892 TTPArgAsnGluGlnPheTTPValIleGlyValSerAlaHisLeuPheAlaVal 911
Db 3002 TGGAGGAACGACAGCTTCTGGGTCAATGGTGGTATTTCTGCGCATTTATTTGCCGCTTC 3061
Qy 912 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 931
Db 3062 CAGGCTCTCTGAAGGTGCTGCTGATCGACACAGAGCTTCACTGTCACCTTAAGGCC 3121
Qy 932 ThrGlyAspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTTPThrThrLeuLeu 951
Db 3122 ACT--GACCAAGAGAGTGTATTTGCCGAGCTCTACATGTTCAAGTGACACACGCTTCTG 3178
Qy 952 IleProThrThrLeuLeuIleIleAsnValIleGlyValValAlaGlyIleSerAsp 971
Db 3179 ATCCACCAACCACTATTTTGATCATCAACCTGGTGGGCTGCTGCGCATTTCTTAC 3238
Qy 972 AlaIleAsnAsnGlyTyrGlnSerTTPGlyProLeuPheGlyLysLeuPheAlaPhe 991
```

```
3239 GCAATCAATAGCGGTTTACCAGTTCATGGGACCTCTTTTTCGGGAAGCTCTTCTTTGGCTTC 3298
Qy 992 TTPValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1011
Db 3299 TGGGTGATTGTCCACTGTACCTCCCTTCCCAAGGGCTCATGGGAAGACGACGACG 3358
Qy 1012 ProThrValValIleTTPSerIleLeuLeuAlaSerIlePheSerLeuLeuTTPVal 1031
Db 3359 CCGACCATTTGCTGTCTGGGTATCTCTTCCGCTCGATCTTTTCCCTGATGTGGTT 3418
Qy 1032 ArgIleAspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIleAsn 1051
Db 3419 CGTATCGATCCATTCACCCCGGCTCACTGGTCTCTGATATCGGAATGTGCATCAAC 3478
Qy 1052 Cys 1052
Db 3479 TGC 3481

RESULT 11
LOCUS AY104236 Zea mays PC0121439 mRNA sequence. 2872 bp mRNA linear HTC 16-OCT-2002
DEFINITION AY104236
VERSION AY104236.1 GI:21207314
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 2872)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 2872)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
source
1..2872
Location/Qualifiers
1..2872
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizDB:637064"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
```

Alignment Scores:
Pred. No.: 0 Length: 2872
Score: 3350.50 Matches: 617
Percent Similarity: 84.84% Conservative: 88
Best Local Similarity: 74.25% Mismatches: 96
Query Match: 59.32% Indels: 30
DB: 8 Gaps: 8

US-10-627-132-30 (1-1052) x AY104236 (1-2872)

QY 246 AlaArgGlnProLeuSerArgIysValSerIleAlaSerSerIysValAsnProTyrArg 265
DB 13 GCACGGCTACCTCTAAGTCGCATAGTTCGGATATCTCCAAACAGCTTAACCTTTATCGG 72
QY 266 MetValIleValValLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHis 285
DB 73 ATCGTGATTGTTCTCCGGCTTATCATCTATGTTCTTCTTCAATCATGATACTCAT 132
QY 286 ProValProAspAlaIleGlyLeuTyrLeuValSerIleIleCysGluIleTyrPheAla 305
DB 133 CCAGTGGAAAGATGCTTATGGTGTGGCTTGTATCTGTTATTTGTGAAGTTGGTTGCC 192
QY 306 IleSerTyrIleLeuAspGlnPheProIysTyrPheProIleAspArgGluTyrIleu 325
DB 193 TTGTCTTGCTCTTAGATCAGTTCCCAAGTGTATCTTATCAACCGTGAACCTTACCTC 252
QY 326 AspArgLeuSerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAla-ValAs 345
DB 253 GATAGACTTGCAATTGAGATATGATAGGAGGGTGGCCATCCAGTTGGGCTCCCAATCGA 312
QY 345 pLeuPheValSerThrValAspProLeuIysGluProProLeuValThrAlaAsnThrVa 365
DB 313 TGTCTTTGTTAGTACAGTGCATCCACTTAAGGAACCTCTCTAATTACTGGCACTGT 372
QY 365 IleSerIleLeuAlaValAspTyrProValAspIysValSerCysTyrValSerAspAs 385
DB 373 CCGTGTCCATTTCTGTGTGGATTACCTGTGACAAAGTATCATGTATGTTCTTCATGA 432
QY 385 pGlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgIysTr 405
DB 433 CGGTTCAGCTATGTTGACTTTTGAAGCGCTATCTGAACCGCAGAGTTTGGCAAGAAATG 492
QY 405 pValProPheCysIysPheGlyIleGluProArgAlaProGluPheTyrPheSerIle 425
DB 493 GGTTCCTTTTGCAGAAACACAAATATTGAACTAGGGCTCCAGAGTTTACTTTGCTCG 552
QY 425 uIysValAspTyrLeuIysAspIysValGlnProThrPheValGlnGluArgAlaMe 445
DB 553 AAAGATAGATTACCTAAAGCAACAATAACAACCTTCTTTGTGAAAGAAAGCGGCTAT 612
QY 445 tIysArgGluTyrGluGluPheIysValArgIleAlaLeuValAlaIysAlaMetIly 465
DB 613 GAAGAGGGAGTGTGAAGATTCAAGTACGGATCGATGCCCTTGTGCAAAAGCGCAAAA 672
QY 465 sValProAlaGluGlyTyrIleMetIysAspGlyThrProTyrProGlyIysAsnThrAr 485
DB 673 AATACCTGAGGAGGGCTGACCATGGCTGATGGCACTCCYTGCCCTGGGAATAACCTAG 732
QY 485 gAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyIysAspThrGluGl 505
DB 733 AGATCATCCMGAAATGATCCAAGTATCTTGGGCCACAGTGGTGGCTTGACACCGATGG 792
QY 505 yAsnGluLeuProArgLeuValTyrValSerArgGluIysArgProGlyPheGlnHisH 525
DB 793 GAATGAGTTGCCACGGCTGTTTATGTTTCTCGTGAAGAGCGCAGGCTCCACACCA 852
QY 525 sIysIysAlaGlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaPr 545
DB 853 CAAGAAGCGTGTGCCATGAATGCTTTGATTGCGGTATCAGCTGTCTTGACGAAATGGTGC 912
QY 545 oPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerIysAlaIleArgGluAl 565
DB 913 TTATCTTCTTAATGTGGATTGTGATCACTACTTCAATAGACCAAAAGCTCTTAGAGAGC 972
QY 565 sMetCysPheLeuMetAspProGlnValGlyArgIysValCysTyrValGlnPheProGl 585
DB 973 TATGTGTTTCATGATGGATCCAGCATAGGAAGGAAACCTTCTGATGTTTCAGTTCCACA 1032
QY 585 nArgPheAspGlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAs 605
DB 1033 AAGATTTGATGATAGACTTCATGATCATGATATGCAAAACCGGAACATTTCTTCTTGA 1092

QY 605 pIleAsnMetIysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysVa 625
DB 1093 TATTAATATGAAGGCTAGATGGCATCTCAGGACCTGTTTATGTGGGAACAGGATGCTG 1152
QY 625 lPheArgArgGlnAlaLeuTyrGlyTyrAsnPro-----ProIysGlyProIysArgPr 643
DB 1153 TTTCAATAGCAGGCTTGTATGGCTATGATCTCTGATTATGACAGAGCTGATTGGAGCC 1212
QY 643 oIysMetValThrCysAspCysCysProCysPheGlyArgIysIysAs- 659
DB 1213 TAAACATTATCAATAAAGTTGCTGTGGC-----GGAAGAAAAAGAGACAGAGCTA 1266
QY 660 -----LysHisAlaIysAspGlyLeuPro----- 667
DB 1267 TATTGATTCCAAAACCGGTGATATGAAGAGAACAGAACTTCCTGGCTCCCATCTTCAACAT 1326
QY 668 -----GluGlyThrAlaAspMetGlyValAspSerAspIysGluMetLeuMe 683
DB 1327 GGAAGATATAGNAGAGGATTGAA-----GGTTACGAGGATGAAGAGTCACTGCTTAT 1380
QY 683 tSerHisMetAsnPheGluIysArgPheGlyGlnSerAlaAlaPheValThrSerThrLe 703
DB 1381 GTCTCAGAAGAGCTTTGGAGAAACGCTTTGGCCAGTCTCCAATTTTATTGTCATCCACCTT 1440
QY 703 uMetGluGluGlyValProProSerSerSerProAlaAlaLeuLeuIysGluAlaI 723
DB 1441 TATGACTCAAGGTGCATACCCCTTCAACAAACCCAGGTTCCCTGCTAAAGAGAACTAT 1500
QY 723 eHisValIleSerCysGlyTyrGluAspIysThrAspTyrGlyLeuGluLeuGlyTyrI 743
DB 1501 ACATGTCATTAGTTGTGATATGAGATAAACAGAAATGGGGAAAGAGATCGGATGGAT 1560
QY 743 eTyrGlySerIleThrGluAspIleLeuThrGlyPheIysMetHisCysAspGlyTyrAr 763
DB 1561 ATATGGCTCTGTACTGAAGATATTTTAACTGGTTTCAAGATGCATGCAAGAGTTGGAT 1620
QY 763 sSerValTyrCysMetProIysArgAlaAlaPheIysGlySerAlaProIleAsnLeuSe 783
DB 1621 ATCCATCTACTGCATGCCACTTCGGCCCTTGGCTTCAAGGGTTCCTGCTCGGATTAATCTTTC 1680
QY 783 rAspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIlePhePheSerAr 803
DB 1681 TGATGCTCTCAACCAAGTGTACGCTGGCTCTTGGTTCAAGTTGAAATTTCTACTTAGCAG 1740
QY 803 gHisSerProLeuLeuTyrGlyTyrIysAsnGlyAsnLeuIysTyrLeuGluArgPheAl 823
DB 1741 ACATGCTCTATCTGGTATGTTAC--AATGGAAGGCTAAAGCTTCTGGAGAGACTGGC 1797
QY 823 aTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyrCysThrLe 843
DB 1798 ATACATCAACACCATTTGTTATCCAATTCATCTATCCCATCTAGTAGCATACTGCGTCT 1857
QY 843 uProAlaValCysIleLeuThrGlyIysPheIleMetProSerIleSerThrPheAlaSe 863
DB 1858 TCCTGCTATCTGTTTACTCACCAAAATTTATTTATCTCGGATGCAATATTATGCTGG 1917
QY 863 rLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTr 883
DB 1918 GCGCTCTTCATCTGCTCTTTTGGTTCATCTTCGCCACTGGTATTTTGGAGCTTCGATG 1977
QY 883 pSerGlyValSerIleGluGluTyrTyrArgAsnGluGlnPheThrValIleGlyIysVa 903
DB 1978 GAGTGGTGTGGCATTTAGAGATTGGTGAGAAATGAGCAGATTTTGGTCAATTTGGTGGCAC 2037
QY 903 lSerAlaHisLeuPheAlaValGlnGlyLeuLeuIysValLeuAlaGlyIleAspTh 923
DB 2038 CTCTGCATCTCTTTGCTGTGTTCCAAAGGTCTCTTAAAGTGTAGCAGGATCGACAC 2097
QY 923 rAsnPheThrValThrSerIysAlaThrGlyAspGluAspAspGluPheAlaGluLeuTy 943
DB 2098 AAACCTTCACGTCATCATCAAGGCACACC--GATGATGATGGTGTGTTTGTGAGCTGTA 2154
QY 943 rAlaPheIysTrpThrThrLeuLeuIleProProThrThrLeuLeuIleIleAsnValI 963

Db 2155 TGTGTTCAAGTGGACAACTCTTCTGATCCCCACCACTGTGCTGTGATTAACCTGGT 2214
 Qy 963 eGlyValValAlaGlyLeuSerAspAlaIleAsnAsnGlyTyrGlnSerTrpGlyProLe 983
 Db 2215 TGGTATATGCTGGAGTGTCTGATGCTATCAACAGTGGCTACCAATCATGGGTCCACT 2274
 Qy 983 uPheGlyValLeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysG1 1003
 Db 2275 ATTGGGAGCTGTTCTTTGCAATCTGGTGATCTCCACCTCTACCTTTCTCGAAGGG 2334
 Qy 1003 YLeuMetGlyArgGlnAsnArgTrpProThrValValValIleTrpSerIleLeuLeuAl 1023
 Db 2335 TCTCATGGGAAGCAGAACCCACACCGACCATCGTCATTGTTGGTCCGCTCTTCTGC 2394
 Qy 1023 aSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThr---LysG1 1042
 Db 2395 TTCCATATTCCTGCTGTGGTGAAGATCGACCCCTTCATATCCCTACCCAGAAAGGC 2454
 Qy 1042 YProAspValArgGlnCysGlyIleAsnCys 1052
 Db 2455 TCTTCCGCTGGCAGTGTGTGTAACCTGC 2485

RESULT 12

CD726831

LOCUS

DEFINITION EST027 Cucurbita pepo testa substracted cDNA Cucurbita pepo cDNA
 clone CES similar to Cellulose synthase, mRNA sequence.

ACCESSION

CD726831

VERSION

CD726831.1

KEYWORDS

EST.

SOURCE

Cucurbita pepo

ORGANISM

Cucurbita pepo

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.

REFERENCE

1. (bases 1 to 1874)

Bezold, T.N., Mathews, D., Loy, J.B. and Minocha, S.C.

Molecular analysis of the hull-less seed trait in pumpkin:

Expression profiles of cell wall related genes during development

Unpublished (2003)

Contact: Subhash Minocha, Dennis Mathews, Brent Loy, Todd Bezold

Dr. Minocha

University of New Hampshire

Rudman Hall, Durham, NH 03824, USA

Tel: 603 862 3840

Fax: 603 862 3784

Email: sminocha@cisunix.unh.edu

Degenerate primers and Taq were used to amplify cDNA for TOPO TA

(Invitrogen, Carlsbad, CA) cloning. Sequencing was performed three

times using the Dyanamic ET Terminator Sequencing kit (Amersham

Pharmacia Biotech Inc, Piscataway, NJ). Sequences were identified

by NCBI BLAST(X).

FEATURES

Location/Qualifiers

1..1874

/organism="Cucurbita pepo"

/mol_type="mRNA"

/db_xref="taxon:3663"

/clone="CES"

/dev_stages="20 days post-anthesis"

/note="Organ: Testa; Total RNA was isolated from 20 day

post-anthesis testa tissue and used in a subtraction

hybridization procedure as according to the Clontech

PCR-Select cDNA Subtraction kit (Prill17-1) (Clontech, Palo

Alto, CA).

ORIGIN

Alignment Scores:

Pred. No.: 6,998-255 Length: 1874

Score: 2525.00 Matches: 467

Percent Similarity: 83.57% Conservative: 57

Best Local Similarity: 74.48% Mismatches: 77

Query Match:

DB:

US-10-627-132-30 (1-1052) x CD726831 (1-1874)

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Query Match:

DB:

US-10-627-132-30 (1-1052) x CD726831 (1-1874)

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Query Match:

DB:

US-10-627-132-30 (1-1052) x CD726831 (1-1874)

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

```
QY 648 -----CysAspCysCysProCysPheGlyArgGlySerGlyHisAlaLys 663
Db 1022 TCTGTCTCTTGGTGGCTGCTCTTGTCTGCTGCTTCCCAAG---AAGATCTCAAAA 1078
QY 664 Asp-----Gly 665
Db 1079 GATCCGACTGAGATTTCAGAGAGATGCAAAAGAGAGAGAGCTTCATGCTGCAATCTTTAC 1138
QY 666 LeuProGluGlyThrAlaAspMetGlyValAlaAspSerAspGlyGluMetLeuMetSerHis 685
Db 1139 CTCAGGAA-----ATAGATAAATATATGATGATGATGATGATGATGATGATGATGATGAT 1192
QY 686 MetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGlu 705
Db 1193 CTGAGCTTTTGAGAAACTTTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1252
QY 706 GluGlyGlyValProSerSerSerProAlaAlaLeuLeuLysGluAlaIleHisVal 725
Db 1253 AATGGCGAGTTTCGGAATCTGCAATCTCTCGACTTTGATCAAGAGCAATTCATGTC 1312
QY 726 IleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeuGlyTrpIleTyrGly 745
Db 1313 ATTAGCTGTGGTTATGAGAGAGATCCACTTCGGGAAAGAGATGCTGCTGATATATGG 1372
QY 746 SerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArgGlyTrpArgSerVal 765
Db 1373 TCAGTGACTGAGGATATCTTAACGGGCTTCAAGATGCTTTCGAGAGGGTGGAGTCCATC 1432
QY 766 TyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArg 785
Db 1433 TACTGATGCCATTATAGGCCAGATTCAAAGGTCGCGACCAATTAACCTTTCTGATGCT 1492
QY 786 LeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePhePheSerArgHisSer 805
Db 1493 CTCACCAAGTTCTTCGATGGGCACTTGGATCTGTTGAGATTTTCTTAGCAGACTGT 1552
QY 806 ProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuLysPheAlaIle 825
Db 1553 CCATTATGTTATGGATTTCGAGCGCGCGCCCTCAATGCTCCAAAGATGCTTACATA 1612
QY 826 AsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyrCysThrLeuProAla 845
Db 1613 AACACCATGTTCTATCCCTTACCTCGCTCCCTCTGTTGCTTACTGCTCATTTGCTGCA 1672
QY 846 ValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThrPheAlaSerLeuPhe 865
Db 1673 ATCTGCTGCTCACAGAAAGTTTCATTCATTCACAGCTCTCGAACCTAGCAAGTACCTT 1732
QY 866 PheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGly 885
Db 1733 TTTCCTGGTCTCTTCTTGTCCATCATTCACGAGTGTCTTCGAGTGTGCTTGGAGTGT 1792
QY 886 ValSerIleGluGluTrpTrpArgAsnGluGlnPheTrpValIleGlyValSerAla 905
Db 1793 GTTAGCATCGAGATATATGGGTACAGAGCAATTCGGTAAATCGGAGCGCTCCGGCA 1852
QY 906 HisLeuPheAlaValGln 912
Db 1853 CATCTCTTTGCGGCTCTTCCA 1873
```

RESULT 13

```
CL978864
LOCUS
DEFINITION
  OsIFCC044639 Oryza sativa Expressed Sequence Tag (EST) clone
  (cDNA) genomic, genomic survey sequence.
ACCESSION
  CL978864
VERSION
  CL978864.1 GI:52412231
KEYWORDS
  GSS.
SOURCE
  Oryza sativa (indica cultivar-group)
  Oryza sativa (indica cultivar-group)
  Eukaryote; Viridiplantae; Streptophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
  Ehrhartoideae; Oryzeae; Oryza.
```

```
REFERENCE
  1 (bases 1 to 3474)
  Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
  Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
  Wang, G. K. S., Deng, X. W. and Wang, J.
  An analysis of transcriptional regulation of the rice genome and
  its comparison to Arabidopsis
  Unpublished (2004)
  Contact: Chen Chen
  Department of Bioinformatics
  Beijing Institute of Genomics
  Chinese Academy of Sciences, Beijing 101300, China
  Tel: 86-10-80481559
  Fax: 86-10-80488676
  Email: chenchen@genomics.org.cn
  Rice genomic sequence.
  Class: exon-trapped.
  Location/Qualifiers
    1..3474
    /organism="Oryza sativa (indica cultivar-group)"
    /mol_type="genomic DNA"
    /db_xref="taxon:39946"
    /clone_lib="Oryza sativa Expressed Library"
    /note="Oryza sativa exon trapped genomic sequences"

ORIGIN
Alignment Scores:
Pred. No.: 7,986-232 Length: 3474
Score: 2311.00 Matches: 489
Percent Similarity: 57.96% Conservative: 177
Best Local Similarity: 42.56% Mismatches: 280
Query Match: 40.92% Indels: 204
DB: 9 Gaps: 28

US-10-627-132-30 (1-1052) x CL978864 (1-3474)

QY 4 SerAlaGlyLeuValAlaGlySerHisAsnArgGluLeuValLeuIleArgGlyHis 23
Db 238 TCCAGCTCGCTCTTCCCGCGGGTTCACAGC-----GTGACACGCGGCAC 285
QY 24 GluAspProLysProLeuArgAlaLeuSerGlyGlnVal-----CysGluIle 39
Db 286 GTGATGGAGAGAGAGCGCTGCTCGCGAGGCGAGCGTGTGCGCTGATGTGCGAGGG 345
QY 40 CysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCys 59
Db 346 TCGCGGTCCAAAGCATCATCGCAACCGGCGCGCGCGCATCTCCCTCTGC---GAGTGC 402
QY 60 GlyPheProValCysArgProCysTyrGluArgGluGlyThrGlnAsnCys 79
Db 403 GACTTCAAGATCTGCTGAGCTGCTTCCAGCGCGCTCAAGGCGCGCGCTGTGTC 462
QY 80 ProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAspAsp 99
Db 463 CCGGGGTGCAAGAGCGCTGACAGCACCGGAGTGGAGAGGTGTGTGCGCGTCC--- 519
QY 100 AspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGln 119
Db 520 -----AACCACGACGCCATCAACAGGCG 543
QY 120 ArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArg 139
Db 544 CTGTCGCTGCGC-----CACGGG--- 561
QY 140 MetSerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIleProPro 159
Db 562 -----CACGGGCGATGGCCCAAGATGAGAGGCGGCTGTGCTGCT--- 599
QY 160 IleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGlyTyr 179
Db 599 ----- 599
QY 180 GlyHisGlyGluValSerSerLeuHisLysArgIleHisProTyrProValSer--- 198
Db ----- 198
```

Db	600	-----GGTGAAGCAGAACGGGGGGGGCCCGCGCGAGTTTCGACCAACAACCGCTGGCTCTTT	653
Qy	199	---GluProGlySerAlaLysTrpAspGluLysGluValSerTrpLysGluArgMet	217
Db	654	CGAGACCAAGGGCACCTACGGCTACGGCAACGCAATCTGGCCGAGGACGACGGCGT---	710
Qy	218	AspAspTrpLysSerLysGlnGlyLleLeuGlyGlyGlyAlaAspProGluAspMetAsp	237
Db	711	-----GGCGGGGCA-----CCCCAAGGAGCTGAT	734
Qy	238	AlaAspValAlaLeuAsnAspGluAlaArg-GlnProLeuSerArgLysValSerIleAl	257
Db	735	-----GAGCAAGGCATCGCGCGCGCTGACCCGCAAGCTCCGGATCCA	776
Qy	257	aSerSerLysValAsnProTrpArgMetValIleValArgLeuValValLeuAlaPh	277
Db	777	GGCGCGCGGTATCAGCCCGTACAGGCTGCTGGTCTCTGATCCGGCTGGTGGCGTGGGTT	836
Qy	277	ePheLeuArgTrpArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValse	297
Db	837	GTTCCTGATGTGGCGCATCAAGCACCAAGCAGGAGCGCCATCTGCTGTGGGGGATGTC	896
Qy	297	rIleIleCysGluIleTrpPhealalleSerTrpIleLeuAspGlnPheProLysTrpPh	317
Db	897	CATCGTGTGGAGCTTGGTTCGCTTGTCTGGTGTGACACGCTGCCACGCTGCCCAAGCTGTG	956
Qy	317	eProIleAspArgGluThrTrpLeuAspArgLeuSerLeuArgTyrGlu-----ArgG	335
Db	957	CCCCATCAACCGCGCCACCGACCTTGACGTGCTCAAGGACAACTTCGAGAGCCCAAGCC	1016
Qy	335	uGlyGluPro-----SerLeuLeuSerAlaValAspLeuPheValSerThrValAs	352
Db	1017	GAGCAACCCACCGGCAAGTCGACCTCCCGGGATCGACATCTCTGCTCCACCGCCGA	1076
Qy	352	pProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAlaValAs	372
Db	1077	CCCGGAGAGGAGCCCTGCTGTGTGTGACGGGCAACACCATCTCTCCATCTCGCCGCCGA	1136
Qy	372	pTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLeuThrPh	392
Db	1137	CTACCCCGTGCACAAAGCTGGCGTCTACGTGTCCGACGCGGGGGCGTCTGTGACGTT	1196
Qy	392	eGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLysLysPh	412
Db	1197	CGAGGCCATGCGCGGAGCGGCGAGCTTCGCCAACCTGTGGTGCCATCTTCGCCGAAGCA	1256
Qy	412	eGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeuLysAs	432
Db	1257	CGAGATCGAGCGGAGGAACCCGACAGCTACTTCAACCTCAAGAGGAGACCGTTCAAGA	1316
Qy	432	plysValGlnProThrPheValGlnGluArgArgAlaMetLysArgGluTyrGluGluPh	452
Db	1317	CAAGGTGAAGGGCGACTTCGTCAAGGACAGCGCGCGGTGAAGCGGAGTACACAGATT	1376
Qy	452	eLysValArgIleAsnAlaLeu-----	459
Db	1377	CAAGTCCCGGTCAATGGCTCCCGACGCCATCCGCGCGCGTCCGACCGGTACCAAGC	1436
Qy	459	-----	459
Db	1437	CGCGGAGGAGATCCAGGCGCATGAACCTGCAGCGGGAGAGATGAAGCCGCGCGCGACGA	1496
Qy	460	----ValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrPr	478
Db	1497	GCAGCAGCTGGAGCGCATAAAGATCCCAAGCGCAGCGTGG---ATGCGCGACGCGCAGCA	1553
Qy	478	oTrpProGly-----AsnAsnThrArgAspHisProGlyMe	490
Db	1554	CTGGCCGGGGGACGTGGCTGCAGCGCTCCCGGAGCACCGGAGGGCGGACCAACCGGGGAT	1613
Qy	490	rIleGlnValPheLeu-----GlyHisSerGlyGlyHis-----	501
Db	1614	CATACAGGTGATGCTAAAGAGCTCCAGTCCAGTCCAGCAGCAGCGCGCGCAGCATGGA	1673

[illegible]

QY 813 ----AsnGly-----AsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrTh 828
 Db 2739 AACAAACAATATCCATTGAGGTTCTCCAGCGCATCTTACTTCAACCTGG 2798
 QY 828 rIeTyrProPheThrSerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLe 848
 Db 2799 CATCTACCCCTTCACTCGCTCTCTCTCATCGCTACTGCTTCTCTCCGGCGCTCTCCCT 2858
 QY 848 uLeuThrGlyLysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIleAl 868
 Db 2859 CTCTCTGGGCGATTCTCATGTCAGACGCTCAACGTCACCTCTCTCACTACCTGCTCAT 2918
 QY 868 leuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIle 888
 Db 2919 CATCACTACACGCTCTGCTCTCTGCGCATCTGGAGATCAAGTGTCTCGGATCGCGCT 2978
 QY 888 eGluGluTrpTrpArgAsnGluGlnPheTrpValIleGlyGlyValSerAlaHisLeuPh 908
 Db 2979 GGAGGAGTGTGGCGAAGCAGCAGTCTTGCTCATCGCGGACACGAGCGCCACCTCGC 3038
 QY 908 eAlaValValGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValTh 928
 Db 3039 CGCGCTCTCTGAGGGCTTCTCAAGGTCTATCGCGCATCGAGATCTCTTCACTCCTCAC 3098
 QY 928 rSerLysAlaThrGlyAspGlu---AspAspGluPheAlaGluLeuTyrAlaPheLysTr 947
 Db 3099 CTCACGACGAGCTCGCGGACGAGTCTGACGACGAGTTCGCGGAGCTCTACGCGTCAAGTG 3158
 QY 947 pThrThrLeuLeuLeuProThrThrThrLeuLeuIleIleAsnValIleGlyValValAl 967
 Db 3159 GAGTCTCTCATGATACCGCGCTCACCATCATATGATCATACCTGTCGCCATCGCGT 3218
 QY 967 aGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLe 987
 Db 3219 CGGCTTCAGCGGACCATCTACAGCACCATCCCGAGTGGAGCAGCTTCTCGTGGAGT 3278
 QY 987 uPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyAr 1007
 Db 3279 CTCTCTCAGCTTCTGGTCTCTCGGCATCTTACCCCTTCGCCAAGGGCTCATGCGGAG 3338
 QY 1007 gGlnAsnArgThrProThrValValIleTrpSerIleLeuLeuAlaSerIlePheSe 1027
 Db 3339 GAGGGGAGGACACCCACCATCTGTCTACGTCTGGTCTGGCTCTGCTGCCATCACCATCTC 3398
 QY 1027 rLeuLeuTrpValArgIleAspPro 1035
 Db 3399 CTCTGCTCTGATTGCAATCAAGCT 3423

RESULT 14

CL964957
 LOCUS OsIPCC011369 Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.
 DEFINITION
 ACCESSION CL964957
 VERSION CL964957.1 GI:52384602
 KEYWORDS GSS.
 SOURCE Oryza sativa (indica cultivar-group)
 ORGANISM Oryza sativa (indica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
 REFERENCE 1 (bases 1 to 2595)
 AUTHORS Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G. K. S., Deng, X. W. and Wang, J.
 TITLE An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis
 JOURNAL Unpublished (2004)
 COMMENT Department of Bioinformatic
 Beijing Institute of Genomics
 Chinese Academy of Sciences, Beijing 101300, China
 Tel: 86-10-80481559

Fax: 86-10-8048676
 Email: chenchen@genomics.org.cn
 Rice genomic sequence.
 Class: exon-trapped.
 FEATURES
 source
 1..2595
 Location/Qualifiers
 /organism="Oryza sativa (indica cultivar-group)"
 /mol_type="genomic DNA"
 /db_xref="taxon:3946"
 /clone_lib="Oryza sativa Express Library"
 /note="Oryza sativa exon trapped genomic sequences"

ORIGIN

Alignment Scores:
 Pred. No.: 2,21e-213 Length: 2595
 Score: 2134.50 Matches: 418
 Percent Similarity: 63.5% Conservative: 127
 Best Local Similarity: 48.7% Mismatches: 261
 Query Match: 37.7% Indels: 51
 DB: 9 Gaps: 13
 US-10-627-132-30 (1-1052) x CL964957 (1-2595)
 QY 232 AspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnProLeuSer 251
 Db 43 GATCCTGAGCCTGCATCTTCAAAAGTGGTGATGACCATGACATACCTGAGCCTCTCAGC 102
 QY 252 ArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleValValArg 271
 Db 103 AGCAGGCTGTCGTACCGTCCGGCAGCTCAACATGATACCGCGCGCGCTCGCGTCCGG 162
 QY 272 LeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAspAlaIle 291
 Db 163 CTCGTCTCATCGCGCGCTTCTTCGCTACCGCGTCAACCGCGCGCTCGCGCGCCAC 222
 QY 292 GlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAsp 311
 Db 223 GCGCTGTGGGAGCGTCCGTGCGGTGAGCTCTGGCTCGCGCGCTCATGGTGTATCGCC 282
 QY 312 GlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSerLeuArg 331
 Db 283 CAGCTCCCGAAGCTCTCCCGCGCCACCGCGTCACTACCTCGACAGGCTGGCCTCAAG 342
 QY 332 TyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSerThrVal 351
 Db 343 TACGAGAAGCGCGGAGCGCTCGCGGTGCGCGCGTGGCGGTGGAGCTGTCTCGTCCGCGCG 402
 QY 352 AspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAlaVal 371
 Db 403 GACGCGCGAGGAGCGCCCGCTGCGGACGCGAACACCGTCTCTCGGTGTCTCGCGCG 462
 QY 372 AspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLeuThr 391
 Db 463 GACTACCGCGCGCGCGCTGCTGCTACGTTCACGACGCGCGCGCGCATGCTGTGTG 522
 QY 392 PheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLysLys 411
 Db 523 TTCGAGTCTCTTTCAGCGCGCGCGCTTTCGAGGCGGTGGATCCCTTCTCGCGCGCG 582
 QY 412 PheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeuLys 431
 Db 583 CACGCGGTGGAGCGCGCGCGCGCTTCTTCTGCGCGCGCGCGCTGCTGCTGCTGCTGCT 642
 QY 432 AspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGluTyrGluGlu 451
 Db 643 GACGCGCGCGCGCGCTTCTGTCAGACCGCGCGCGCATGAGAGGAGTACGAGGAG 702
 QY 452 PheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGluGlyTrp 471
 Db 703 TTCAGGTGGAGTGAACCACTCTGCGCGGAGGCGCGCAAGGTGCGCGGAGGAGATGG 762
 QY 472 IleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGlyMetIle 491

Db 763 ATCATGTGCGACGCGCGTGGCGGGGAAACAATCGCGGGATCAACCTCGCCATGATT 822
 Qy 492 GlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeu 511
 Db 823 CAGGTGCTCTGGGTCAACCGCGGACGCGACGTCACGCGCGGAGCTCCCGCGCTC 882
 Qy 512 ValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGlyAlaMet 531
 Db 883 TTCTACGTGTGCGGGAGAGACGACCGGGTTCCGGACACACGCGCAAGCGCGGCCATG 942
 Qy 532 AsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAsp 551
 Db 943 AACCGCTGTCTACGGGTATCGCGGTCTTACCAATGGCGCTACGTCTCAACCTGGAC 1002
 Qy 552 CysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeuMetAsp 571
 Db 1003 TGCACCACTCGTCAACAACAGACGCGCTGAGAGCGGATGTCTTTATGATGGAC 1062
 Qy 572 ProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIleAsp 591
 Db 1063 CCGTTCGCGGGAACAGACGTGTCTGTCAGTTCCGACTGCGCGACAGCGGC----- 1116
 Qy 592 ValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLysGlyLeu 611
 Db 1117 -----GGCGCGCGCGCGCGGCGACTCCGCTCTTCTCGATATCGAGATCGAATGCTC 1170
 Qy 612 AspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeu 631
 Db 1171 GATGGCATCCAGGCGCGGTGTACGTGCGGTTCCGGTTGCTGCTTTCAGCGGAGCGGCTG 1230
 Qy 632 TyrGlyTyrAsnPro-----ProLysGlyProLysArgProLysMetValThr 647
 Db 1231 TACGGTTCGAGCCCGCGCGCGCGCTGACAGCGCGACGACATGGATACGCGCGCGGAC 1290
 Qy 648 CysAspCysProCysPheGlyArgLysLysArgLysHisAla----- 662
 Db 1291 TGGAGGAGGATGTCTGCTTCGGGAGGGGAGAGGATGAACGCGATGAGGAGTCCATG 1350
 Qy 663 -----LysAspGlyLeuProGluGlyThrAlaAspMetGly 674
 Db 1351 TCCGCGGTGTGCTGCTGACTCGGAGGATGACTCCGACGACACAGGAGGAGGAGGCG 1410
 Qy 675 ValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGln 694
 Db 1411 CGGGGAGGAGGAGGAGCTGCGCGCTACCGCGCGCGCTGGAGCGCCACTTCGCGCAG 1470
 Qy 695 SerAlaAlaPheValThrSerThrLeuMetGlu-----GluGly 707
 Db 1471 TCTCCCGGTTTCATCGCTCGCGGTTTCGAGGAGCAAGGCGCGGCGGCGGCGACGCT 1530
 Qy 708 GlyValProProSerSer-----SerProAla-----AlaLeuLeuLysGluAlaIleHisVal 725
 Db 1531 GGCTCGCTGACGCGACCGTGGCGCGACGCGATCTCTCTCAAGAGGCGCATCATGTC 1590
 Qy 726 IleSerCysGlyTyrGluAspLysThrAspTyrGlyLeuGluLeuGlyTyrIleTyrGly 745
 Db 1591 GTTACTGCGGTTTCAGAGAGGAGGACAGATGGGCGCAAGAGATTGGATGATGTACGGC 1650
 Qy 746 SerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArgGlyTyrArgSerVal 765
 Db 1651 GGC-----GGTGTGCGACGCGGTTCCGATGTCACGCGCGGGTGGTGTGCGCG 1701
 Qy 766 TyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArg 785
 Db 1702 TACTGCTCGCGCGAGGCGCGGTTACGGCGGTACGCGCGCGCGCGCGCGGACGTC 1761
 Qy 786 LeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIlePhePheSer---ArgHis 804
 Db 1762 CTCGGGGGGCGCTCGCGGCGAGCGGTGGCGGCATGGGATCTCTGAGCGCGCGGAC 1821
 Qy 805 SerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTyrLeuGluArgPheAlaTyr 824
 Db 1822 AGCCCCGCTGCGCGCGCGCGCTCGGG-----CTCTGACGCGGCTGGGCTAC 1872

Qy 825 IleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyrCysThrLeuPro 844
 Db 1873 GTGGCGCGCGCGCTACCGCTCGCTCCCTCCCGCTCACCGCTACTTGGCGCTCCCG 1932
 Qy 845 AlaValCysLeuLeuThrGlyLysPheIleMetProSer---IleSerThrPheAlaSer 863
 Db 1933 GCGGTCTGCTCTCACCAGGCAAGTCCACGTTCCCGGACGACGTAGCTACTACGCGC 1992
 Qy 864 LeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTyr 883
 Db 1993 GTCTCTCTCATCTGCTCTTCTCGTGGCGCGCTCGGTGCGCTCGAGCTGAGGTGG 2052
 Qy 884 SerGlyValSerIleGluGluTyrTyrArgAsnGluGlnPheTyrValIleGlyVal 903
 Db 2053 AGCGGTGCTCATCGCGCGGTGTGGGAGAGAGAGAGCTCTGGATGGTCACGCCACA 2112
 Qy 904 SerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAlaGlyIleAspThr 923
 Db 2113 TCGCGAGCGCTCGCGCGGTGTCCAGGCGATCTCAGCGGTGTCACCGGATCGACGTC 2172
 Qy 924 AsnPheThrValThrSerLysAlaThr----- 932
 Db 2173 GCCTTCTCCAGGAGCGCGAGCTTCGCCACCAAGCGACCGCGCGGAGAAACGATGAC 2232
 Qy 933 GlyAspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTyrThrThrLeuLeuIle 952
 Db 2233 GGGAGAGAGAGCGCATTTGGCATCGGAGATTACGATCGGTGGTACGACCTTGTGGTG 2292
 Qy 953 ProProThrThrLeuLeuIleAsnValIleGlyValValAlaGlyIleSerAspAla 972
 Db 2293 GCCCGAGCGAGCTGGTGGTGGGACCTCTCGCGGGGTGTAGCGGGGTGGCGGTACGGG 2352
 Qy 973 IleAsnAsnGly---TyrGlnSerTyrGlyProLeuPheGlyLysLeuPhePheAlaPhe 991
 Db 2353 GTGACCAACGAGTACTACCACTGTGGGGCGCGCTCGCGCGAGCTGGCGCTCGCGCGG 2412
 Qy 992 TrrValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1011
 Db 2413 TGGTGTGGCGGACCTTCAGGATTCCTCGGGGGTCTCTGGCGCGCGAGACCGGCGG 2472
 Qy 1012 ---ProThrValValIleTrrSerIleLeuLeuAlaSerIlePheSerLeuLeuTrr 1030
 Db 2473 CCGCGGACCATCGCGGTCTCTGTCGTGTCTTCTCTCGGTGGGTCTGCTCTCTG 2532
 Qy 1031 ValArgIleAspProPheIleValArgThrLysGlyProAspValArgGln 1047
 Db 2533 GTCCACCGCGCTCTTCTCGCGCGCGACGCGCGCGCCACGACGAGGACAA 2583

RESULT 15

CL974686

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CL974686 3048 bp DNA linear GSS 21-SEP-2004
 OsIFCC026364 Oryza sativa Express Library Oryza sativa (indica
 cultivar-group) genomic, genomic survey sequence.

CL974686
 CL974686.1 GI:52403885

GSS.
 Oryza sativa (indica cultivar-group)

Oryza sativa (indica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzeae; Oryza.

1 (bases 1 to 3048)

Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,

Zhao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,

Wong, G. K. S., Deng, X. W. and Wang, J.

An analysis of transcriptional regulation of the rice genome and

its comparison to Arabidopsis

Unpublished (2004)

Contact: Chen Chen

Department of Bioinformatic

Beijing Institute of Genomics

Chinese Academy of Sciences, Beijing 101300, China

Tel: 86-10-80481559
 Fax: 86-10-80488676
 Email: chenchen@genomics.org.cn
 Rice genomic sequence.
 Class: exon-trapped.
 Location/Qualifiers
 1. 3048
 /organism="Oryza sativa (indica cultivar-group)"
 /mol_type="genomic DNA"
 /db_xref="taxon:39946"
 /clone_lib="Oryza sativa Express Library"
 /note="Oryza sativa exon trapped genomic sequences"
 ORIGIN
 Alignment Scores:
 Pred. No.: 3 77e-197 Length: 3048
 Score: 1982.00 Matches: 439
 Percent Similarity: 55.29% Conservative: 157
 Best Local Similarity: 40.72% Mismatches: 281
 Query Match: 35.09% Indels: 201
 DB: 9 Gaps: 24
 US-10-627-132-30 (1-1052) x CL1974686 (1-3048)
 QY 75 GlyThrGlnAsnCysProGlnCysLysThrArgTyrLysArgLeuLys----- 90
 DB 15 GGGGAAGAGGGGATCCGGAACCGCGCGCGTCCGCGCGCGCGCGCGCGCGCGG 74
 QY 91 GlySerProArgValAlaGlyAspAspAspGluGluAspIleAspAspLeuGluHisGlu 110
 DB 75 CGGGCTCGCGCGCGCGCGCGCGCGCGGTGAAGTTCGGCGCGCGCGCGCGCGCGG--- 131
 QY 111 PheAsnIleAspAspGluAsnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGln 130
 DB 132 -----GCG 134
 QY 131 IleThrGluAlaMetLeuHisGlyArgMetSerTyrGlyArg----- 144
 DB 135 GTACGTGACCTGTCCGCGGAGGACATCGACATGAGGCGGAGCTGGCGGCGGACTACAC 194
 QY 145 -----GlyProAspAspGlyAspGlyAsnAsnThrProGlnIleProIlelle 161
 DB 195 CAATACACGGTGCAGATCCCGCGAGCGCGCGACAA-----CCAGCCGAT 239
 QY 162 ThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrGlyHis 181
 DB 240 GCTGAACGCGCGGAGCGCGCGCGTGTGGCGAT-----GAAGCGCGGAGGACA 287
 QY 182 GlyGluValSerSerSerLeuHisLysArgIle-----HisProTyrProValSer 198
 DB 288 GTACGTGTCCAACTCCCTCTTACCGGCGGCTTCAACAGCGCCACCGCGCCCACTCAT 347
 QY 199 GluProGly----- 201
 DB 348 GGACAAGGTATCGAGTGTGTCTCCACCGCGAGATGGCGCGCGCGCAAGGGTTCGG 407
 QY 202 -----SerAlaLysTrpAspGluLysLysGluValSerTrpLysGluArg----- 216
 DB 408 CTGCGCATGCCCGCTGCGAGCGGAGCCATGCGCAACGAGCGCGCGGAGAGCTCGA 467
 QY 217 ---MetAspAspTrpLysSerLysGln-GlyIleLeuGlyGlyGlyValAspProGluAs 235
 DB 468 ACCCTGCGAACCTTGGCGGCGAGCGGCGGCGCGCGCGCGCGCGCGCTCCCGCGAG-- 525
 QY 235 pMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSe 255
 DB 526 -----CAGNAGCGGTTCAGCGCGCTCAGCGCGCTCAGCGCGAGATCCC 560
 QY 255 rIleAlaSerSerLysValAsnProTyrArgMetValIleValValArgLeuValValSe 275
 DB 561 GATCGCGAGTCCGTTGATCAGCGCGCTACCGGATCTTCATCGTGTATCCCGATGCTGCT 620
 QY 275 uAlaPheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLe 295

DB 621 CCTCTTCTACCTCAGTGGCGCATCCGCAACCAACATGGAGGCGCTCTGCTGTGGGG 680
 QY 295 uValSerIlelleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProly 315
 DB 681 GATGTCATCGTCTGGAGCTCTGTTGGCTTCTCATGGCTGCTCGACATGCTCCCAA 740
 QY 315 sTrpPheProIleAspArgGluThrTyrIleuAspArgIleuSerLeuArgGluG1 335
 DB 741 GGTGAACCGGTCAACCGGAGCACCGCTCCCGTGTCTCAAGGAGAAGTTCGAGACGCC 800
 QY 335 uGlyGluPro-----SerLeuLeuSerAlaValAspLeuPheValSerth 350
 DB 801 GTGCGCGTCCAAACCGCGCGCGCTCCGACCTCCCGGCGCTTCGACGTGTTGTTCTCCAC 860
 QY 350 rValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAl 370
 DB 861 CGCGCATCGGAGAGAGCGCGTCTCACACGGCGGACCACTCTCTCATCTCCGCG 920
 QY 370 aValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLe 390
 DB 921 CGTGACTACCGCGTGCAGAAAGCTCGCTGTCTACGTTCGACGACGCGCGCGCTGCT 980
 QY 390 uThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysly 410
 DB 981 CACCTTCGAGGCGCATGGCGGAGCGCGGAGCTTCGCCAACGTTGGTGGTCCCTTCTGCAA 1040
 QY 410 sLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrIle 430
 DB 1041 GAAGCACCATCGACCCCGCGCAACCCCGGACAGCTACTTTCAGCTCAGGCGGACCCGAC 1100
 QY 430 uLysAspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGluTyrG1 450
 DB 1101 GAAGGCAAGCGCGCGCAACGACTTCGTCAAGGACCGCGCGAGGGTGAAGCGGAGTTCGA 1160
 QY 450 uGluPheLysValArgIleAsnAlaLeu----- 459
 DB 1161 CGAGTTCAAGGTGCGCATCAATGGCTCCCGGACTCCATCCGCGCGCGCTCCGACGCGTT 1220
 QY 460 -----ValAl 461
 DB 1221 CAACGCGCGGAGGACATGAAGATGCTCAAGCACCTCCGCGAGACGGCGCGCGCGCTC 1280
 QY 461 aLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrProTrpProG1 481
 DB 1281 GGAGCAGCCCAAGGTGAAGAGGCGGACGCTGG--ATGGCGGACGCTCGCACTGGCGCGG 1337
 QY 481 Y-----AsnAsnThrArgAspHisProGlyMetIleGluVa 493
 DB 1338 GACATGGCGCGGCTGCGCGCGCGCGACCGCAAGGGCAACCGCGCGGATCCTCCAGT 1397
 QY 493 lPheLeuGlyHisSer-----GlyGlyHisAspThrGlu----- 504
 DB 1398 GATGCTCAAGCGCGCGTCCCGCGCGCGCTCTACGGGATGACGACGAGCACCATGAT 1457
 QY 505 -----GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysAr 519
 DB 1458 CGACTTCAGCGAGCTCGACATCCGCTCCCGATGCTCTACATGTCCGCGGAGAGCG 1517
 QY 519 pProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSerAl 539
 DB 1518 CCGCGGCTTACGACCAACAAAGAGCGCGCGCGCATGAACCGCTCTGCTCCGATGCTCCGC 1577
 QY 539 aValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSe 559
 DB 1578 GGTATGTCAAGCGCGCGCTTCATGCTCACTTCGACTCGACCCACTACATCAACAAGCG 1637
 QY 559 rLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysValCy 579
 DB 1638 GCAGCGGTGCGCGAGGCGCATGTCTTCTTCATGGAC---CGCGCGCGCGGAGCGCATCGC 1694
 QY 579 sTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsnAr 599

```
Db      1695 CTACATCCAGTTCCGCGAGCGCTTCGAGGGGATCGACCCCTCCGACCGCTACGCCAACAA 1754
Qy      599 gAsnThrValPhePheAspIleAenMetLysGlyLeuAspGlyIleGlnGlyProValTy 619
Db      1755 CAACACCGTCTTCTTCGATGGCAACATCGCGCGCTCGAGCGCTCCAGGGCCCCATGTA 1814
Qy      619 rValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrglyTyraenProProLysG1 639
Db      1815 CGTCGGCAGCGCTGCGATGTCGCGCTTCGCGCTCTAGCGCTTCAGACCTCCGCCAC 1874
Qy      639 yProLysArgProLysMetValThrCysAspCysCysProCysPheGlyArgLysLysAr 659
Db      1875 CGCCGATACACGGGTGGCTC-----TTCAAGAAGAAAGGT 1913
Qy      659 gLysHisAlaLysAspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAspLy 679
Db      1914 GACCACGTTCAAGACCGGAGTCCGACACCCAGACGCTCAAGCGCGGAGGACTTCGACGC 1973
Qy      679 sGluMetLeuMetSerHisMetAenPheGluLuLysArgPheGlyGlnSerAlaAlaPheVa 699
Db      1974 CGAG---CTCAGTCGCACCTC---GTCCCCCGCGGTTCGGCAACTCGTCGCCGTTTCA 2027
Qy      699 lThrSerThrLeuMetGluGluGlyGlyValProPro---SerSerSerProAlaAlaLe 718
Db      2028 GCGCTCGATCCCGTCGCGAGTTCAGGCGCGGCCCTCGCCGACCAACCGCGGGTGT 2087
Qy      718 u-----CGAGCGCGCGCTCACCGTCCCGCGCGCGCGTGGACCGCGCGCGGT 2147
Db      2088 GCAGCGCAGCGCGTCCGCGCGCGCTCACCGTCCCGCGCGCGCGTGGACCGCGCGCGGT 2147
Qy      719 uLysGluAlaIleHisValIleSerCysGlyTyrgluAspLysThrAspTTPGlyLeuG1 739
Db      2148 GCGGAGCGCGTTCGCGGTGATCTCGTGTGTAGCAGGACGAGGAGTGGGGCGGACCG 2207
Qy      739 uLeuGlyTTPileTyrglySerIleThrGluAspIleLeuThrGlyPheLysMetHisCy 759
Db      2208 CGTGGGTGGATCTACGGCTCGGTCCAGGAGCGTGGTACCGCGCTACCGGATGCACAA 2267
Qy      759 sArgGlyTTPArgSerValTyrglyCysMetProLysArgAlaAlaPheLysGlySerAlaPr 779
Db      2268 CCGCGGTGGCGCTCCGTCTACTGCATCACCAGCGCGACGCGTTCCTGGGCGCGCGCC 2327
Qy      779 rIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTTPAlaLeuGlySerValGluI1 799
Db      2328 GATCAACTCACCGACAGCTCCACAGTGTCTCCGTGGGCGCGCGCTCCGTTCGAT 2387
Qy      799 ePhePheSerArgHisSerProLeuLeuTyrglyTyrglyAsnGlyAsnLeuLysTTPLe 819
Db      2388 CTTCCTCTCCCGAACACGCGTTCCTGGCGTGGCGAAG-----CTCATGCTGCT 2438
Qy      819 uGluArgPheAlaTyrlleAsnThrIleTyrgProPheThrSerLeuProLeuLeuAl 839
Db      2439 GCAGCGGATCTCCTACCTCAACGTCGGCATCTACCCCTTCACCTCCATCTTCCTCCTCGT 2498
Qy      839 aTyrgCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSe 859
Db      2499 CTACTGCTTTCATCCCGCGGCTCTCCCTCTTCGCGGCTTTCATCGTCCAGAGCTCGA 2558
Qy      859 rThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLe 879
Db      2559 CATCGCTTCTTGTACTCTCTCACCATGACCATCACCCTCGTCCGCTCGGCATCCT 2618
Qy      879 uGluMetArgTTPSerGlyValSerIleGluTTPTrpArgAsnGluGlnPheTTPVa 899
Db      2619 CGAG-----GGGTTCGTGAAGTGTATGCG 2622
Qy      899 lIleGlyGlyValSerAlaHisLeuPheAlaValGlnGlyLeuLeuLysValLeuAl 919
Db      2623 -----GGGTTCGTGAAGTGTATGCG 2642
Qy      919 aGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGlu---AspAspGl 938
Db      2643 CGGGATAGATCTCCTTCACGCTGACGGGAAGCGCGCGCGGCGGACGACACAGGAGCAT 2702
```

Search completed: March 12, 2005, 14:26:04
Job time : 5268 secs

```
Qy      938 uPheAlaGluLeuTyrtAlaPheLysTTPThrThrLeuLeuIleProProThrThrLeuLe 958
Db      2703 CTACGGCGACCTTACATCGTCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2762
Qy      958 uIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrg1 978
Db      2763 GATGGTGAACATCATCGCCATCGCCTTCGCTTCGGGAGGACCATCTACAGCGCAACACC 2822
Qy      978 nSerTTPGlyProLeuPheGlyLysLeuPhePheAlaPheTTPValIleValHisLeuTy 998
Db      2823 GCGGTGGGGAAGTTTCATCGCGCGCGCTTCTTCAGCTTCTGGGTGCTCGCCACCTCAA 2882
Qy      998 rProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIleTr 1018
Db      2883 CCCCTTCGCCAAGGGCTCATGGGGCGGCGCGGCAAGACCCACCATCGTCTTCGTCGTG 2942
Qy      1018 pSerIleLeuLeuAlaSerIlePheSerLeuLeuTTPValArgIleAspPro 1035
Db      2943 GTCGGGCTCTCTCAATCACCGTCTCCCTGCTCTGGTGGTGGTGGTGGTGGTGGTGGTGG 2994
```

Qy	1861	CGAGCTGCGACGACCGATACGCTAAACAGGAAACACCGTCTTCTTCGACATCAACATGAAGGG	1920
Db	1861	CGAGCTGCACGACCGATACGCTAAACAGGAAACACCGTCTTCTTCGACATCAACATGAAGGG	1920
Qy	1921	GCTCGACGGCATCCAAGSCCGGTGTAGCTCGGACAGGGTGCCTGTTCGGGCGCCAGGC	1980
Db	1921	GCTCGACGGCATCCAAGSCCGGTGTAGCTCGGACAGGGTGCCTGTTCGGGCGCCAGGC	1980
Qy	1981	GCTCTACGGCTACAACCTCCTCCAAAGGAGCCCAAGAGGCCCAAGATGTGTACCTTCGCACTG	2040
Db	1981	GCTCTACGGCTACAACCTCCTCCAAAGGAGCCCAAGAGGCCCAAGATGTGTACCTTCGCACTG	2040
Qy	2041	CTGCCCCGTGTTTCGGCCGCAAGAAAGCGGAAACAACCGCAAGGACGGGCTGCGGAGGCGAC	2100
Db	2041	CTGCCCCGTGTTTCGGCCGCAAGAAAGCGGAAACAACCGCAAGGACGGGCTGCGGAGGCGAC	2100
Qy	2101	CGCTGATATGGGAGTACGATACGACACAGGAGATGCTCATGTCCACATGAACATTCGAGAA	2160
Db	2101	CGCTGATATGGGAGTACGATACGACACAGGAGATGCTCATGTCCACATGAACATTCGAGAA	2160
Qy	2161	GCGETTCGGGCAGTCCGCGCGTTCGTCAAGTCGACGCTGATGAGGAAAGCGCGCTCCC	2220
Db	2161	GCGETTCGGGCAGTCCGCGCGTTCGTCAAGTCGACGCTGATGAGGAAAGCGCGCTCCC	2220
Qy	2221	TCCTTCGTCGAGCCCCCGCGCTCCTCAAGAGGGCCATCCATGTCACTACGTTCGGCTA	2280
Db	2221	TCCTTCGTCGAGCCCCCGCGCTCCTCAAGAGGGCCATCCATGTCACTACGTTCGGCTA	2280
Qy	2281	CGAGGACAAAGCCGACTTGGGGGCTGGAGCTGGGGTGGATCTACGGGTGATACGAGGGA	2340
Db	2281	CGAGGACAAAGCCGACTTGGGGGCTGGAGCTGGGGTGGATCTACGGGTGATACGAGGGA	2340
Qy	2341	CATCCTGACGGGGTTCAAAGATGCACTGCGCGGGTGGCGTCCGTGTACTGCAATCCGAA	2400
Db	2341	CATCCTGACGGGGTTCAAAGATGCACTGCGCGGGTGGCGTCCGTGTACTGCAATCCGAA	2400
Qy	2401	CGCGGCGCGGTTCAAGGGTTCGCGCGATCAATCTATCGGACCGTCTCAACAGAGTGCT	2460
Db	2401	CGCGGCGCGGTTCAAGGGTTCGCGCGATCAATCTATCGGACCGTCTCAACAGAGTGCT	2460
Qy	2461	CCGCTGGGCGCTGGGGTCCGTCGAGATCTTCTCAGCGGGCAGAGCCCCCTGCTGACGG	2520
Db	2461	CCGCTGGGCGCTGGGGTCCGTCGAGATCTTCTCAGCGGGCAGAGCCCCCTGCTGACGG	2520
Qy	2521	CTACAAGAACGGCAACTCAAGTGGCTGGAGCGCTTCGCGCTACATCAACACCACTCATCTA	2580
Db	2521	CTACAAGAACGGCAACTCAAGTGGCTGGAGCGCTTCGCGCTACATCAACACCACTCATCTA	2580
Qy	2581	CCCCCTCACCTCGCTCCCGCTGCTCGCTACTGCAACCTCCCGCGCTGTGCCTCTCTCAC	2640
Db	2581	CCCCCTCACCTCGCTCCCGCTGCTCGCTACTGCAACCTCCCGCGCTGTGCCTCTCTCAC	2640
Qy	2641	CGGCAAGTTTCATATGCGGTGATTAGACAAGTTTCGCCAGCGCTCTTCTCATTCGCGCTCTT	2700
Db	2641	CGGCAAGTTTCATATGCGGTGATTAGACAAGTTTCGCCAGCGCTCTTCTTCTCATTCGCGCTCTT	2700
Qy	2701	CATGTCCATCTTCGCGACGGGCACTCCGAGATGCGGTGGAGCGGGGTGAGCATCGAGGA	2760
Db	2701	CATGTCCATCTTCGCGACGGGCACTCCGAGATGCGGTGGAGCGGGGTGAGCATCGAGGA	2760
Qy	2761	GTGTTGAGGAAACGAGCAGTCTTCGGGTCAATCGGCGGCGGTGTCCGCGCATCTCTTCGCGCT	2820
Db	2761	GTGTTGAGGAAACGAGCAGTCTTCGGGTCAATCGGCGGCGGTGTCCGCGCATCTCTTCGCGCT	2820
Qy	2821	CGTGCAGGGCTGTCTAAGGTCTTCGCGCGGGATCGAACAACTTCAACCTCACTCTCCAA	2880
Db	2821	CGTGCAGGGCTGTCTAAGGTCTTCGCGCGGGATCGAACAACTTCAACCTCACTCTCCAA	2880
Qy	2881	GGCCACCGGCGACGAGGACGAGTTCGCGAGCTCTAGCCCTTCAAGTGGACCAAGCT	2940
Db	2881	GGCCACCGGCGACGAGGACGAGTTCGCGAGCTCTAGCCCTTCAAGTGGACCAAGCT	2940
Qy	2941	CCTCATCCCCGCCACCAACGCTGTCTCATTTAAAGTTCATCGGGGTGTGGCGCGGCACTCTC	3000

[illegible]

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Hotta, Y., Iida, J., Iida, Y., Ikeda, R., Inamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakahama, Y., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

TITLE
Collection, mapping, and annotation of 28K full-length cDNA clones from japonica rice

JOURNAL
Unpublished
3 (bases 1 to 3631)

REFERENCE

Kikuchi, S.
Direct Submission
Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression: 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

COMMENT
This clone is one of the 32K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Iehikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Yamamoto, M. and Nakahama, Y.

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Iehibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Oca, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES
source

1. 3631
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clones="J023081B08"

ORIGIN

Query Match 78.2%; Score 2691.2; DB 8; Length 3631;
Best Local Similarity 89.6%; Pred. No. 0;
Matches 2953; Conservative 0; Mismatches 323; Indels 18; Gaps 5;

Qy	67	CCTAGCACGCCGCCCTCCCGCCGATGAGGCGCAGGCGCGGGCTGGTGGCCCGGCTCGCA	126
Db	71	CGTTGCGCGCGCTTGTGTGCGCGCCATGGAGGCGAGCGCGGGCTGGTGGCCCGGCTCGCA	130
Qy	127	CRACCCGACGAGCTGGTGTGATCCGGGCGCAGAGGACCCCAAGCCGCTGGCGGCGCT	186
Db	131	CAACCCGACGAGCTGGTGTGATCCGGGCGCAGAGGACCCCAAGCCGCTGGCGGCGCT	190
Qy	187	GAGCGGCGAGGTGTGCGAGATATGCGGCGCAGAGGTCGGGCTCACGGTGACGCGCGACCT	246
Db	191	GAGCGGCGAGGTGTGCGAGATATGCGGCGCAGAGGTCGGGCTCACGGTGACGCGCGACCT	250
Qy	247	CTTTCGTCGCTGCAACGAGTCGGCTTCCCGGTGTGCGCGCCCTGTCTACGATACGAGCG	306
Db	251	CTTTCGTCGCTGCAACGAGTCGGCTTCCCGGTGTGCGCGCCCTGTCTACGATACGAGCG	310
Qy	307	CCGGGAGGGCAGCAGAACCTGCCCGCAGTGCAGAGCGCTACAAGCGCTCAAGGGGAG	366
Db	311	CCGGGAGGGCAGCAGAACCTGCCCGCAGTGCAGAGCGCTACAAGCGCTCAAGGGGAG	370
Qy	367	CCCGAGGGTTGCCGGGGACGATGACGAGGAGGACATCGACGACCTGGAGCACGAGTTCAA	426
Db	371	CCCGAGGGTTGCCGGGGACGATGACGAGGAGGACATCGACGACCTGGAGCACGAGTTCAA	430
Qy	427	CATCGACGACGAGAATCAGCAGAGGAGGAGCTGGAGGGCAA-----CATGCAGAACAGCCA	480
Db	431	CATCGACGACGAGAATCAGCAGAGGAGGAGCTGGAGGGCAA-----CATGCAGAACAGCCA	490
Qy	481	GATCACGGAGGCGATCTGCACGGCAGGATGAGTACGGGAGGGGCCCCACGACGGCGA	540
Db	491	CATCACGGAGGCGATCTGCACGGCAGGATGAGTACGGGAGGGGCCCCACGACGGCGA	550
Qy	541	CGGCAACAAACACCCCGCAGATCCCGCCATCATCACGGCTCCCGTCCCGTGGCGGTAG	600
Db	551	CGGCAACAGACACCCCGC---TCCCGCGATCATCACGGGCGCTCGTCCGTCGGGTAG	607
Qy	601	CGGTGAGTTTCCGATTACCAACGGGTATGCGCACGGGAGGTTCGTTCTTCCTGCAAA	660
Db	608	CGGGGAGTTTCCGATATCGAACAGCCATGCGCCATGCGAGTTTCTCTCTTCTTCTGCAAA	667
Qy	661	GGCATCCATCCCTGCTGTCGTCGACGAGGAGTGCACAGTGGGAGGAGGAGGAAGA	720
Db	668	GGCATCCATCCCTGCTGTCGTCGACGAGGAGTGCACAGTGGGAGGAGGAGGAAGA	727
Qy	721	AGTGAGCTGGAAGGAGGATGGAACGACGACTGGAAGTCCAAAGCAGGGCATCTCGCGCGG	780
Db	728	GGTGAGCTGGAAGGAGGATGGAACGACGACTGGAATCCAAAGCAGGGCATCTCGCGCGG	787
Qy	781	CG---CCGATCCCGAAGACATGAGCGCGGACGTGGGACCTGAACGACGAGGCGGAGGCC	837
Db	788	CGCCCCCGATCCCGACGACTACGACGCGCGACGCTCCCACTGAACGACGAGGCGGAGGCC	847
Qy	838	GCTGTGAGGAAGGTGTGATCCGTCGACCAAGGTGAACCCGTACCGGATGGTGTGATCGT	897
Db	848	GCTGTGAGGAAGGTGTGATCCGTCGACCAAGGTGAACCCGTACCGGATGGTGTGATCAT	907
Qy	898	GGTGGCTCTCGTGTGCTCGCCCTTCTCTCTCCGGTACCGGTATCTGTGACCCCGTCCCGGA	957
Db	908	CCTCCGCTCTCGTGTGCTCGCCCTTCTCTCTCCGGTACCGGTATCTGTGACCCCGTCCCGGA	967
Qy	958	CGCCATCGGGCTGTGGCTGTCTTCATCATCTGCGAGATCTGGTTGCGCATCTCTCGAT	1017
Db	968	CGCCATCGGGCTGTGGCTGTCTTCATCATCTGCGAGATCTGGTTGCGCATCTCTCGAT	1027
Qy	1018	CCTCGACCAAGTTCCTCAAGTGTTCCTCCACCGGAGACCTGCTGACCGCCCTCTC	1077
Db	1028	CCTCGACCAAGTTCCTCAAGTGTTCCTCCACCGGAGACCTGCTGACCGCCCTCTC	1087
Qy	1078	CCTCAGGTACGAGGAGGAGGCGGTCGCTGCTGTCGGCGGTGAGACCTGTTGTCGAG	1137
Db	1088	CCTCAGGTACGAGGAGGAGGCGGTCGCTGCTGTCGGCGGTGAGACCTGTTGTCGAG	1147

		this category is not included in IRGSP standard"
gene	25510..27251	
misc_feature	/gene="P0418B08.5"	
	25510..27251	
	/gene="P0418B08.5"	
	/notes="probably inactive due to including stop codon(s) in CDS	
	pseudogene, transposase"	
	/pseudo	
gene	complement(29005..36438)	
mRNA	/gene="P0418B08.6"	
	complement(join(<29005..29109,29221..29352,29652..29818,30448..30637,31719..31850,32464..32571,32702..32797,32891..32998,33106..33278,34215..34590,35076..35219,35314..35988,36106..36171,36271..36438))	
	/gene="P0418B08.6"	
	/notes="start and end point are not identified"	
CDS	complement(join(29005..29109,29221..29352,29652..29818,30448..30637,31719..31850,32464..32571,32702..32797,32891..32998,33106..33278,34215..34590,35076..35219,35314..35988,36106..36171,36271..36438))	
	/gene="P0418B08.6"	
	/notes="similar to Oryza sativa chromosome 7, OSJNBa061L20.19"	
	/codon_start=1	
	/product="hypothetical protein"	
	/protein_id="BAD33414.1"	
	/db_xref="GI:50725886"	
	/translation="MASDTSSDCCSSKNDKVAQSDVANSISEEETSDSDKYI TKDLVSTLKLDFSGSIKRRARTKPVRRQKQTKLEYDQDKSGSAESPKRSLN ITFYSNIELSNEQSRIENSAFGLLNFQRCALPLSEVKWIASHTDVSCSDIVVG RSLPINTNTFLTGPLNGGLRIKNDNDAGRHQFHGSKTFLISFFGKLLSDKGV NKLSEDLVRCFVVALSTFLGCPNSDTPSPKYLEPLSDIKSSKWNMSKFVYEWLMT YIAKFQESKSEKQTSKTLGCGHLAIINLDFQGIIRNIPIGPPIRSYVKGGMK EYKIDCKTGDGKPLDDLSVCKWFTREHSDFTYKQAQLAAKVLDCVMSKAS LFTQDQTTTIREKSEEEKSGSGKGGNNVAHGNAPPLEQNHNTDVSQLPKTP QTKAEHTKSNINNTISKGAPTDKKAASDFVCSQQSNNSVASTRAKRLHSHS PLVSGISPSKYSSDKRKARVKOPLKATEPINTQPDREESKDALFVEPCTIPA KKEVOPTKLNJESSSEFVIDIEGPYDAEDITGHTTDTKTKFLVNVSSBEHNSQD TODESDINPKNSQPKAPLHHFDDKTKNSKASLGHLFSOTSGTGNMIDIEHSMQ SSOSSPDIWNPSPIAOMEPHQHQAERQYSMIRIIDLNASANCSGTBNHLYRPK RIVHPILMDNVKVTWSSLSSKLSLSPRGVDTYVLNAYAKKIANDQNNKNEVNFYFFH RYSLIPLFDYDNHWFVITVDIIGHHFVLDSYDENNKYHKKIQGLLPGFTAMWEEF SYVENKFSFDIQYPPITRQNG"	
gene	40242..41638	
mRNA	/gene="P0418B08.7"	
	join(<40242..40308,40655..40900,40996..41112,41239..41339,41570..41638)	
	/gene="P0418B08.7"	
	/notes="start and end point are not identified"	
CDS	join(40242..40308,40655..40900,40996..41112,41239..41339,41570..41638)	
	/gene="P0418B08.7"	
	/notes="predicted by GeneMark.hmm etc."	
	/codon_start=1	
	/product="hypothetical protein"	
	/protein_id="BAD33415.1"	
	/db_xref="GI:50725887"	
	/translation="MPATSELEATATWELDEAGKAWGGALPHPPHPKPTAAVATRR	
Query Match	48.1%;	Score 1657.8; DB 8; Length 165909;
Best Local Similarity	69.7%;	Pred. No. 8.4e-209;
Matches 2954;	Conservative 0;	Mismatches 322; Indels 963; Gaps 14;
Qy	67	CCTAGCACGCGCGCTCCGCGCGCATGAGGCGAGCGCGGCTGGTGGCCGCGCTCGCA 126
Dd	5926	CGATGCGGCGCGTGTGTGCGCGCCATGAGGCGAGCGCGGCTGGTGGCCGCGCTCGCA 5985
Qy	127	CAACCGGAACGAGCTGGTGTGATCCGGGSCACAGGAGCCCAAGCGCTCGCGGCGCT 186
Dd	5986	CAACCGGAACGAGCTGGTGTGATCCGGGSCACAGGAGCCCAAGCGCTCGCGGCGCT 6045
Qy	187	GAGCGGCGAGGTGTGGAGATATGCGCGCACAGGTCGGGCTCA CGGTGGAGCGGCACT 246

Db	7123	TACCGGATGGTGATCATCTCTCCGTCCTCGTGGTGTCTCGGCTTTCTTCTCCCGTACCGCATC	7182
Qy	941	CTGCACCCCGTCCCGAGCGCATCGGCTGTGGCTCGTCTCCATCATCTGCGAGATCTGG	1000
Db	7183	CTCACCGGTGCCGAGCGCATCCGCTGTGGTCACTCTCCATCATCTGCGAGATCTGG	7242
Qy	1001	TTGCGCATCTCTCGATCTCTCGACCAAGTTCCTCCCAAGTGGTTCCTCCATCGACCGCGAGACG	1060
Db	7243	TTGCGCGTGTCTGGATCTCTCGACCAAGTTCCTCCCAAGTGGTACCGCATCGACGAGGAGAC	7302
Qy	1061	TACCTTCGACCGCTCTCCCTCAGGTACGAGAGGGAAGGGAGCGGTCGCTGCTGTCCGGC	1120
Db	7303	TACCTTCGACCGCTCTCCCTCAGGTACGAGAGGGAAGGGAGCGGTCGCTGCTGTCCGGC	7362
Qy	1121	GTGACCTGTTCGTGAGCAGCGTGGACCCGCTCAAGGAGCGCGCTGTGTGACCGCAAC	1180
Db	7363	GTGACCTGTTCGTGAGCAGCGTGGATCCGCTCAAGGAGCGCGCTGTGTGACCGCAAC	7422
Qy	1181	ACCGTGTCTCCATCTCTCGCGTGAATACATCCCGTGGACAAAGTGTCTCTGTACGTCTCC	1240
Db	7423	ACCGTGTCTCCATCTCTCGCGTGAATACATCCCGTGGACAAAGTGTCTCTGTACGTCTCC	7482
Qy	1241	GACGACGCGGTGATGCTGACGTTGAGTCTGAGTCTGTCGAGAGCGCGGAGTTCGCGGC	1300
Db	7483	GACGACGCGGTGATGCTGACGTTGAGTCTGAGTCTGTCGAGAGCGCGGAGTTCGCGGC	7542
Qy	1301	AAGTGGGTGCTCTCTGCAAGAGTTCGGCATCGAGCCCGCGCGAGTTCATCTTC	1360
Db	7543	AAGTGGGTGCTCTCTGCAAGAGTTCAGCATCGAGCCCGCGCGAGTTCATCTTC	7602
Qy	1361	TCGCTCAAGGTGCTACTCTCAAGGACAAAGTTCAGGCTGAGGAGGAGCGCGC	1420
Db	7603	TCCGAGAGGTGCTACTCTCAAGGACAAAGTTCATCCCACTTCGTCGAGAGCGCGCGC	7662
Qy	1421	GCCATGA-----	1427
Db	7663	GCCATGAAGTGAATATCGATCGCATCGTCACTGCTGATTTCTGTAGAGCTTGACGAAGA	7722
Qy	1428	-----AGAGAGATATGAGGAGTTCAAGG	1451
Db	7723	GCTAACTGTTGTGGGGCATTTGGCATTTGATCGAGCAGAGAGTACGAGGATTTCAAGG	7782
Qy	1452	TCCGGATCAACGCGCTGTGTGCAAGGCTATGAAGTTCGCGGACAGAGGGTGTGATATGA	1511
Db	7783	TGAGGATCAACGCGCTGTGTGCGAAGCGCAGAAAGTTCGCGGCGAAGGGTGGATCATGA	7842
Qy	1512	AGGACGACGCGCTGTGCGCGGGAACACACCCCGGACCAACCCCGGCGATGATCCAGGTGT	1571
Db	7843	AGGACGACGCGCTGTGCGCGGGAACACACCCCGGACCAACCCCGGCGATGATCCAGGTGT	7902
Qy	1572	TCTTGGGCGCACAGCGGGCGCACACACCGAGGGCAACGAGCTGCCCGGCTCGTGTACG	1631
Db	7903	TCTTGGGCGCACAGCGGGCGCACACACCGAGGGCAACGAGCTGCCCGGCTCGTGTACG	7962
Qy	1632	TCTCCCGTGAAGGCGCGGGATTCAGACACCAAGAGGCGCGGCGCCATGAACGCTC	1691
Db	7963	TCTCCCGTGAAGGCGCGGGCTTCAGACACCAAGAGGCGCGGCGCCATGAACGCCC	8022
Qy	1692	T-----	1692
Db	8023	TCGTAAGCGCACCGCCACCATCTCTTTCAATTTTCAAAAACAACAACATCTGCTCGATTTGAT	8082
Qy	1693	-----GATTTCGCTCTCCGCGTGTGACCAACCGCGCCATTC	1729
Db	8083	CGAAATTTCTGGTTTGGCATTTCTGTGTGTGCGCGGTGTGACGACGCGCGCTTC	8142
Qy	1730	ATGCTCAACTTTGGATGTGATCACTATCAATCAACAACGAGGCGCATCCCGGAGGCCATG	1789
Db	8143	ATGCTCAACTTTGGATGTGATCACTATCAATCAACAACGAGGCGCATCCCGGAGGCCATG	8202
Qy	1790	TGCTTCCTCATGAGCCCTCAGGTGCGGAGGAGTCTGCTACGTTCAAGTTCCTCCGCGAGG	1849
Db	8203	TGCTTCCTCATGAGTCCGAGGTGCGGAGGAGGTTTGTACGTGCGAGTTCCCGCGAGG	8262
Qy	1850	TTTCGCGCATCGACGTGCACGACCGATACGCTAACAGGAACACCGCTCTTCTTCGAC---	1906
Db	8263	TTTCGCGCATCGACGTGCACGACCGATACGCTAACAGGAACACCGCTCTTCTTCGAGCTG	8322
Qy	1907	-----	1906
Db	8323	AGCTCTCTCCACACAACTAGATGAATATATAAATCTTTGAAAAATCCAGAGCAAAATCA	8382
Qy	1907	-----A	1907
Db	8383	CGATCGATCGAGCAATGTGACTGAATTTTGTGTGTGCGTCTTGTGTGAGATCATCAGA	8442
Qy	1908	TCACATGAAGGGGCTGCGACGCGCATCCAAAGGCGCGGTGTACGTGCGGACAGGGTGCCTGT	1967
Db	8443	TCACATGAAGGGGCTTGTATGGGATCCAGGGCCGCGTGTACGTGCGGACAGGGTGCCTGT	8502
Qy	1968	TCCGGCGCAGGCGCTCTACGGCTACAAACCTCCCAAGGACCCCAAGAGCCCAAGATGG	2027
Db	8503	TCAGGGCGGAGGCGCTGTACGGATACAAACCCCAAGGACCCCAAGAGCCCAAGATGG	8562
Qy	2028	TGACCTGCGACTCTCGCGTGTCTCGCGCGCAAGAGCGGAAACACCGCAAGAGCGGCG	2087
Db	8563	TGACCTGCGACTCTGCGCTTGTCTCGGAGGAGAGCGGAAAGCGGACAGGACGCGC	8622
Qy	2088	TGCGGAGGCGCACCGC-----	2103
Db	8623	TCCCGGAGGCGCTCGCGCGCGAGCTGCCAAATTCAGAGCTAAGAAGAAAT	8682
Qy	2104	-----TGATATGGGA	2113
Db	8683	ATGGTGTGAGAGATTTTCAGCTGATGGAATAATGAAGTTTGTGTGTGGATCAGGG	8742
Qy	2114	GTAGATAGCACAAGAGATGCTCATGTCCCATATGAATTCAGAGAGCGGTTTCGGGCGAG	2173
Db	8743	ATGAGACGCGACAAGGAGATGCTCATGTCCGATGAATTCGAGAGAGCGGTTTCGGGCGAG	8802
Qy	2174	TCCCGCGCGTTCCTGCTACGCTGATGAGAGGAGGCGGCTCCCTCTTCTGCTCGAGC	2233
Db	8803	TCCCGCGCGTTCCTGCTACGCTGATGAGAGGAGGCGGCTCCCGCGCTGCTCCAGC	8862
Qy	2234	CCCGCGCGCTCTCTAAGAGGCGCATCATGTATCATGCTGCGGCTACGAGGACAGAGC	2293
Db	8863	CCCGCGCGCTCTCTAAGAGGCGCATCATGTATCATGCTGCGGCTACGAGGACAGAGC	8922
Qy	2294	GACTGGGGGCTGG-----	2306
Db	8923	GACTGGGCTCTCGAGGTAAACCGTTATCAGACACTAAGCCATATATTACATTGAGTGAG	8982
Qy	2307	-----AGCTGGGGTGGATCTACGGGTCGATC	2332
Db	8983	TTTGTGTGTGAACGCCATGATGTGTGATGACGCTGGGGTGGATCTACGGGTCGATC	9042
Qy	2333	ACGAGGACATCTCTGACGGGTTCAAGATGCACTCGCGGGGTGGGCTCCGTGTACTGC	2392
Db	9043	ACGAGGACATCTCTAAGGGGTTCAAGATGCACTCGCGGGGTGGAGGTTCGGTGTACTGC	9102
Qy	2393	ATCCGAGCGGCGGCTTCAAGGGGTGGCGCGCATCAATCTATCGGACCGCTCTCAAC	2452
Db	9103	ATCCGAGGAGGCGGCTTCAAGGGGTGAGCGCGCATCAACCTATCTGACCGCTCTCAAC	9162
Qy	2453	CAGGTGCTCGGTGGGCGCTGGGGTCCGTGAGATCTTCTTTCAGCGGCGACAGCCCCCTG	2512
Db	9163	CAGGTGCTCGGTGGGCGCTGGGGTCCGTGAGATCTTCTTTCAGCGGCGACAGCCCCCTG	9222
Qy	2513	CTGTACGGCTACAAGAACGCGCAACCTCAAGTGTGAGGCGCTTCGCTATCATCAACACC	2572
Db	9223	CTGTACGGCTACAAGAACGCGCAACCTCAAGTGTGAGGCGCTTCCTATCATCAACACC	9282
Qy	2573	ACCATCTACCCCTTACCTCGCTCCGCTGCTGCTACTGACCGCTCCCGCGGCTGCG	2632
Db	9283	ACCATCTACCCCTTACCTCTCTCCCTCTCTCGCTACTGACCGCTTACCGCGGCTGCG	9342

CDS	7723.. .81241) /gene="OJ1740_D06.2" /note="supported by full-length cDNA(s): AK065475" complement(join(2772..3329,3566..4732,5547..5616, 7723..7784)) /gene="OJ1740_D06.2" /note="contains EST(s): AU092173(C11159),D22463(C11159) contains full-length cDNA(s): AK065475" /codon_start=1 /product="putative RPT2" /protein_id="BAD33627.1" /db_xref="GI:50726106" /translation="MKFMKSGKDPADFQSGDGVRYVISDLATDVIHVSEVKFYLHK FPLSKSKLVKATIECTDEHIDFGPGVTAFEICAKFCYGMVVTLSPEHNVVA ARCAAELEMTEDVKNGLIKFIDVFNSSIRSWKDSIIVLQSTKALLPSEELKVI GRCDATASVDPANVTWYSYHSRKGMSCTEIVSTGRTSIAPKDWVEDICELDV DLYKRVAVSKSGMSPELLGEALKAYAVRWLPDSVDALVAEDYMRNQCLVETIIV LPLSDKTSGCCRFLLKLLKVAIIIVGAGQHVKEELMRISFQHLKASVKOILLPAAP SGDAHVKLVHNLVORFVARTAMSHNGFVBSDDKMIELNFEHESTIALGELYDGYL SEVADPDLSTAVELATVPEARPVHDSYSAVDAYLKEHPNI SKADKKIKCGLI DVKLSITDASHATQNDRLPURLVVQVLFPOOLRAGSGNALALTDGGHTCAKPIMKD QSDICERIRIPRHNSLANKQATSLSAREVEHFKSEHGGGNSFKDQLGGFLQSRSR IFDKWSKGGKGGKSGTSSQSPPLSAKPADVPSPLPPLPNRRYSVS" /complement(join(10055..10068,13432..13638,13766..13889)) /gene="OJ1740_D06.3" complement(join(10055..10068,13432..13638,13766..13889)) /gene="OJ1740_D06.3" /note="hypothetical ORF predicted by GENSCAN this category is not included in IRGSP standard" 14864..15751 /gene="OJ1740_D06.4" 14864..15751 /gene="OJ1740_D06.4" /note="probably inactive due to no initiation codon in CDS pseudogene, putative HGWP repeat containing protein" /pseudo 16561..17118 /gene="OJ1740_D06.5" 16561..17118 /gene="OJ1740_D06.5" /note="probably inactive due to 5' exon missing in CDS pseudogene, putative HGWP repeat containing protein" /pseudo complement(17924..18233) /gene="OJ1740_D06.6" complement(join(<17924..17955,18104..>18233)) /gene="OJ1740_D06.6" /note="start and end point are not identified" complement(join(17924..17955,18104..18233)) /gene="OJ1740_D06.6" /note="similar to Oryza sativa chromosome 1, P0480E02.21" /codon_start=1 /product="hypothetical protein" /protein_id="BAD33628.1" /db_xref="GI:50726107" /translation="MKMKIIVDDKRNHDEDDVDPPANSVATNDNDKAHQHLHN ASLGELMK" join(18381..18410,18539..18577,19649..19732) /gene="OJ1740_D06.7" join(18381..18410,18539..18577,19649..19732) /gene="OJ1740_D06.7" /note="hypothetical ORF predicted by GlimmerM this category is not included in IRGSP standard" 25476..26366 /gene="OJ1740_D06.8" 25476..26366 /gene="OJ1740_D06.8" /note="probably inactive due to 5' exon missing in CDS pseudogene, TNP2-like transposon protein" /pseudo 26747..28873
CDS	7723.. .81241) /gene="OJ1740_D06.2" /note="supported by full-length cDNA(s): AK065475" complement(join(2772..3329,3566..4732,5547..5616, 7723..7784)) /gene="OJ1740_D06.2" /note="contains EST(s): AU092173(C11159),D22463(C11159) contains full-length cDNA(s): AK065475" /codon_start=1 /product="putative RPT2" /protein_id="BAD33627.1" /db_xref="GI:50726106" /translation="MKFMKSGKDPADFQSGDGVRYVISDLATDVIHVSEVKFYLHK FPLSKSKLVKATIECTDEHIDFGPGVTAFEICAKFCYGMVVTLSPEHNVVA ARCAAELEMTEDVKNGLIKFIDVFNSSIRSWKDSIIVLQSTKALLPSEELKVI GRCDATASVDPANVTWYSYHSRKGMSCTEIVSTGRTSIAPKDWVEDICELDV DLYKRVAVSKSGMSPELLGEALKAYAVRWLPDSVDALVAEDYMRNQCLVETIIV LPLSDKTSGCCRFLLKLLKVAIIIVGAGQHVKEELMRISFQHLKASVKOILLPAAP SGDAHVKLVHNLVORFVARTAMSHNGFVBSDDKMIELNFEHESTIALGELYDGYL SEVADPDLSTAVELATVPEARPVHDSYSAVDAYLKEHPNI SKADKKIKCGLI DVKLSITDASHATQNDRLPURLVVQVLFPOOLRAGSGNALALTDGGHTCAKPIMKD QSDICERIRIPRHNSLANKQATSLSAREVEHFKSEHGGGNSFKDQLGGFLQSRSR IFDKWSKGGKGGKSGTSSQSPPLSAKPADVPSPLPPLPNRRYSVS" /complement(join(10055..10068,13432..13638,13766..13889)) /gene="OJ1740_D06.3" complement(join(10055..10068,13432..13638,13766..13889)) /gene="OJ1740_D06.3" /note="hypothetical ORF predicted by GENSCAN this category is not included in IRGSP standard" 14864..15751 /gene="OJ1740_D06.4" 14864..15751 /gene="OJ1740_D06.4" /note="probably inactive due to no initiation codon in CDS pseudogene, putative HGWP repeat containing protein" /pseudo 16561..17118 /gene="OJ1740_D06.5" 16561..17118 /gene="OJ1740_D06.5" /note="probably inactive due to 5' exon missing in CDS pseudogene, putative HGWP repeat containing protein" /pseudo complement(17924..18233) /gene="OJ1740_D06.6" complement(join(<17924..17955,18104..>18233)) /gene="OJ1740_D06.6" /note="start and end point are not identified" complement(join(17924..17955,18104..18233)) /gene="OJ1740_D06.6" /note="similar to Oryza sativa chromosome 1, P0480E02.21" /codon_start=1 /product="hypothetical protein" /protein_id="BAD33628.1" /db_xref="GI:50726107" /translation="MKMKIIVDDKRNHDEDDVDPPANSVATNDNDKAHQHLHN ASLGELMK" join(18381..18410,18539..18577,19649..19732) /gene="OJ1740_D06.7" join(18381..18410,18539..18577,19649..19732) /gene="OJ1740_D06.7" /note="hypothetical ORF predicted by GlimmerM this category is not included in IRGSP standard" 25476..26366 /gene="OJ1740_D06.8" 25476..26366 /gene="OJ1740_D06.8" /note="probably inactive due to 5' exon missing in CDS pseudogene, TNP2-like transposon protein" /pseudo 26747..28873
CDS	7723.. .81241) /gene="OJ1740_D06.2" /note="supported by full-length cDNA(s): AK065475" complement(join(2772..3329,3566..4732,5547..5616, 7723..7784)) /gene="OJ1740_D06.2" /note="contains EST(s): AU092173(C11159),D22463(C11159) contains full-length cDNA(s): AK065475" /codon_start=1 /product="putative RPT2" /protein_id="BAD33627.1" /db_xref="GI:50726106" /translation="MKFMKSGKDPADFQSGDGVRYVISDLATDVIHVSEVKFYLHK FPLSKSKLVKATIECTDEHIDFGPGVTAFEICAKFCYGMVVTLSPEHNVVA ARCAAELEMTEDVKNGLIKFIDVFNSSIRSWKDSIIVLQSTKALLPSEELKVI GRCDATASVDPANVTWYSYHSRKGMSCTEIVSTGRTSIAPKDWVEDICELDV DLYKRVAVSKSGMSPELLGEALKAYAVRWLPDSVDALVAEDYMRNQCLVETIIV LPLSDKTSGCCRFLLKLLKVAIIIVGAGQHVKEELMRISFQHLKASVKOILLPAAP SGDAHVKLVHNLVORFVARTAMSHNGFVBSDDKMIELNFEHESTIALGELYDGYL SEVADPDLSTAVELATVPEARPVHDSYSAVDAYLKEHPNI SKADKKIKCGLI DVKLSITDASHATQNDRLPURLVVQVLFPOOLRAGSGNALALTDGGHTCAKPIMKD QSDICERIRIPRHNSLANKQATSLSAREVEHFKSEHGGGNSFKDQLGGFLQSRSR IFDKWSKGGKGGKSGTSSQSPPLSAKPADVPSPLPPLPNRRYSVS" /complement(join(10055..10068,13432..13638,13766..13889)) /gene="OJ1740_D06.3" complement(join(10055..10068,13432..13638,13766..13889)) /gene="OJ1740_D06.3" /note="hypothetical ORF predicted by GENSCAN this category is not included in IRGSP standard" 14864..15751 /gene="OJ1740_D06.4" 14864..15751 /gene="OJ1740_D06.4" /note="probably inactive due to no initiation codon in CDS pseudogene, putative HGWP repeat containing protein" /pseudo 16561..17118 /gene="OJ1740_D06.5" 16561..17118 /gene="OJ1740_D06.5" /note="probably inactive due to 5' exon missing in CDS pseudogene, putative HGWP repeat containing protein" /pseudo complement(17924..18233) /gene="OJ1740_D06.6" complement(join(<17924..17955,18104..>18233)) /gene="OJ1740_D06.6" /note="start and end point are not identified" complement(join(17924..17955,18104..18233)) /gene="OJ1740_D06.6" /note="similar to Oryza sativa chromosome 1, P0480E02.21" /codon_start=1 /product="hypothetical protein" /protein_id="BAD33628.1" /db_xref="GI:50726107" /translation="MKMKIIVDDKRNHDEDDVDPPANSVATNDNDKAHQHLHN ASLGELMK" join(18381..18410,18539..18577,19649..19732) /gene="OJ1

QY	1892	ACGCTCTTCTTCGACATCAACATGAAGGGCTCGACGGCATCAAGCCCGGTGTACGTC	1951
Db	1776	ACTGTTTCTTCGATATAACATGAAGGCTAGATGAATTCAGGGTCCAGTGTATGTG	1835
QY	1952	GGACAGGGTGCCTGTTCCGGCCGAGGCGCTCTACGGCTACAACTCCCAAGGACCC	2011
Db	1836	GGCACAGGATGCGTTTTCAAAGGCAAGCTTTGTATGGCTATGACCTCCCAAGGATCCA	1895
QY	2012	AAGAGGCCAAGATGCTGACCTCGACTGCTGCCGCTGCTTCGGCCGCAAGAGCGGAA	2071
Db	1896	AAGGCCCAAGATGAACCTCGACTGCTGCCCATGTTTGGAGCTGCCTCAAAAGAG	1955
QY	2072	CAGCCAAAGGACGGGCTGCCGAGGCGCACCGCTGATATGGAGTAGATGCAAGGAG	2131
Db	1956	AATGCTAAGACTGGTGCAGTTGTAGAAGGAATGGATA-----ATAATGACAAGGAG	2006
QY	2132	ATGCTCATGCTCCACATGAATTCGAGAGAGCGGTTCCGGCAGTCCGGGGGTTCTGCACG	2191
Db	2007	CTGTTGATGTCACATGAATTTTGAAGAAAGTTTGGACAATCAGCAATTTTCGTAAC	2066
QY	2192	TCGACGCTGATGAGGAAGGCGCTCCCTCTTCGTTCGAGCCCGCGCGCTCCCTCAAG	2251
Db	2067	TCAACTTTAATGGAAGAGTGTGTACCTCTTCCTCGAGTCCGGAGCTCTGCTAAAG	2126
QY	2252	GAGGCCATCCATGTCATGCTCGGCTACGAGGACAAAGCCGACTGGGGGCTGGAGCTG	2311
Db	2127	GAAGCCATCCATGTCATGCTGATGATGATGATGAAGACAAATCTGAATGGGACTCGAGCTG	2186
QY	2312	GGGTGATCTACGGGTGATGATGAGGAGCATCTGACGGGGTTCAGATGCACTGCCG	2371
Db	2187	GGGTGATTTACGGTTTCGATCAGCGAGGATATCTGACAGGTTTAAAGATGCAATTCGT	2246
QY	2372	GGGTGGCGTCCGTTACTGCTCATGCCGAGCGGCGCGTTCAGGGGCTCGGCGCGATC	2431
Db	2247	GGCTGAGGTCTAATTACTGTATGCCAAGAGAGCTGCATTTAAGGGTTCAGCTCCATC	2306
QY	2432	AATCTATCGGACCGTCTCAACACGAGTCTCCGTTGGGCGCTCGGCTCGAGATCTTC	2491
Db	2307	AATCTATCAGATCGGCTAAACCAAGTCTCCGCTGGGCTCTTGATCTGTTGAAATTTT	2366
QY	2492	TTGAGCGGACACGCCCCCTGCTGATCGGCTCAAGACGGAACCTCAAGTGGCTGGAG	2551
Db	2367	TTGAGTGGTCACAGCCCTTAATGGTATGGCTAACAAGAAAGTAAGTCAAGTGGCTCGAG	2426
QY	2552	CGCTTGGCTTACATCAACACACATCTACCCCTTCACCTCGCTCCCGCTGCTCGGCTAC	2611
Db	2427	AGGTTTGGGTATGTGAACACATATCTACCCCTTACCTCTTAGCACTCGTTGATAC	2486
QY	2612	TGCACCTCTCCCGCGCTGCTGCTCCTCACCGCAAGTTTCATCATATGCGGTGATTAACG	2671
Db	2487	TGTTGCTCCTCGCATCTGCTGCTTACTGTATAAATTTATCATGCGCGAGATAAGCACC	2546
QY	2672	TTGCGAGCTCTTCTTCATGCGGCTCTTCATGTCATCTTCGCGAGGCGATCCTGGAG	2731
Db	2547	TTTGGCAAGTCTTTCTTCATGCTGCTGTTTGTCTCAATCTTTTCCAGCGGCATCTTGAG	2606
QY	2732	ATCGGTGGAGCGGGTGACATCGAGGAGTGTGGAGGAACGAGAGTTCCTGGGTATC	2791
Db	2607	CTCAGATGGAGCGGAGTAAGCATTTGAGGAATGTTGAGAAACGACGAATTTCTGGGTTATA	2666
QY	2792	GGCGGCTGTCGCGCATCTCTTTCGCGCTGTCGAGGCGCTGCTCAAGTCTCTCGCGGG	2851
Db	2667	GGTGGTGTGCTGCTCACCTCTTGTGTTGTTGCCAAGTCTTCTGAAAGTTTTCAGAGGT	2726
QY	2852	ATGCACACCAATCTTACCGTCACTTCAAGGCCACCGGCGACGAGACGAGTTCGCC	2911
Db	2727	ATTCGACTAAATCTTCACTGTGCATCAAGGGCTAC-----AGACGATGACGATTTTGA	2780
QY	2912	GAGCTCTACGCTTCAAGTGGACACAGCTCTCTCATCCCGGCCACACGCTGCTCATCAT	2971
Db	2781	GAGCTTATGCTTTTAAATGGACAAACCTGCTTATCCCTTCAACCATATCTTAATCATC	2840
QY	2972	AACGTCATCGGCTCGTCCCGGCTATCTCGAGCGCATCAACAAACGGGTACAGTCTCTGG	3031
Db	2841	AACCTTGTGGAGTGTGTTCTCGAGTCTCAGATGCATAAACAAATGGGTACCAAGTCTATGG	2900
QY	3032	GGGCCCCCTTCGGCAAGCTCTTTCGCTTCTGGGTCATGTCACCTCTACCGGTTTC	3091
Db	2901	GGACCTCTAATTCGGGAAGCTCTTCTTTCGCTTCTGGGTGATTTGTCATCTCTACCAATTC	2960
QY	3092	CTCAAGGGGCTCATGGGCGCCAGAACAGGAGCCGCCACCGTTGTTGTCATCTGTCATTC	3151
Db	2961	CTCAAGGCTAATTCGGGAGGCAAAACAGACACCGACTATTGTTGTTATATGTCAGTG	3020
QY	3152	CTGCTGGCTCTCATCTTCTCCTGCTCTGGGTCAGGATCGACCTTTTCATCTGTCAGGACC	3211
Db	3021	CTCTGGCTTCCATCTTCTCTTGGTTCGGTCCGATTCATCTTGTGATGAAACCC	3080
QY	3212	AGGGCCCGGAGCTCAGGAGTGTGGCATCAATTCCTGA	3250
Db	3081	AGGGGACCTTGACACCAAGCAATGTGGACTCAACTGTTAA	3119
RESULT 6			
AY372244			
LOCUS	AY372244	3470 bp	mRNA linear PLN 20-JAN-2004
DEFINITION	Zea mays cellulose synthase catalytic subunit 10 (CesA10) mRNA, complete cds.		
ACCESSION	AY372244		
VERSION	AY372244.1	GI:38532099	
KEYWORDS	Zea mays		
SOURCE	Zea mays		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
REFERENCE	1 (bases 1 to 3470)		
AUTHORS	Dhugga,K.S., Barreiro,R., Appenzeller,L., Wang,H., Niu,X., Carrigan,L. and Tomes,D.		
TITLE	Cellulose formation and its role in determining stalk strength in maize		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 3470)		
AUTHORS	Dhugga,K.S.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-AUG-2003) Trait and Technology, Pioneer Hi-Bred International, Inc., a DuPont company, 7300 NW 62nd Avenue, Johnston, IA 50131, USA		
FEATURES	Location/Qualifiers		
source	1..3470		
	/organism="Zea mays"		
	/mol_type="mRNA"		
	/db_xref="taxon:4577"		
gene	1..3470		
	/gene="CesA10"		
CDS	29..3265		
	/gene="CesA10"		
	/note="CESA10; ZmCESA10"		
	/codon_start=1		
	/product="cellulose synthase catalytic subunit 10"		
	/protein_id="AA223310.1"		
	/db_xref="GI:38532100"		
	/translation="MDAGSVTGLAAGSHMRDELHVMRAREBPNAKVRADVKTRVC ADEVTREDDQPFVACAECPVPCPYEYSESGTCCPCQCNTRYKROKQRCRVBG EBEQEMDDDFDEFPKAPKPEHPVAFDVYSENGEHPAQKWTGQTLSSTFTGSVAG KDLAERMEGSMWKDRIDKWTKEKRLNHDSDDDDDKNEDEYMLLAARAPL WKVPITPSMINPYRIVILRLVLCFLKPLTTPATDAVPLASVLCELWFASFWS ILDQLPKAPVTRTYLDRALRYDREGACRLSPIDFFVSTVDPLKEPPIITANTVL SILAVDYPVDRVSCYVSDGASMLLFDALSETAFARWRVPFCKEFAVFPRAFEFVLS QKIDYLDKQVPTFKVRRAMREYEFKVRINALVAKQAKPEEGVMQDQFPWFGN NTRDPGMTQVYLGNGALDVEGHELPLVYVSRKPGYNHKKAGAMNALVRVSAY LTNAPFLNLDCDHYVNNSKAVREAMCFMDQLGKLCYQFPQSFQDGLDRHRYAN RNVPFDINMKGLDQGVVYVYGCVFNRQALYGYDPPRPEKRPWMTCDWPSWCC CCGGKGKAKDKDGGGEPRLGLLGYRKSQDKLGGSVAGSKSGGLYKK HORAFEEIEEGLEGYDELESSLMSQSKSEKRFQSQSPVFIASTLIVDEGGLPQGA		

DPALIKAEIHLVISCVEEKEIEMWKEIGWLVSVTEIDLTGFKMHCWKSVYCTPR
PAFKGSPINLSDELHVLRLWALGVSVEIFMSRCHPLRYAIGRLKWLREPAYNTYVI
PFTSIPULACTTIPAVCLLTKGFIPTLNNLAIWFIALFSLIATISVLSRWGVSIV
EDWRNEQFVWIGVSAHLFAVFGQFLKVLGVDSFTVTSKAAGDADAFDGLYLFK
WTLILVPTTLIIINMVGIVAGVSDAVNNGVSGWGLFGLKLFPSFWIVHLVPLKGL
MQRNRTPTIIVLWLSILLASIFSLVWVRIDPFIPKAGPILKPGVBC"

ORIGIN

Query Match 45.4%; Score 1563.4; DB 8; Length 3470;
Best Local Similarity 70.9%; Pred. No. 2.8e-136;
Matches 2369; Conservative 0; Mismatches 671; Indels 303; Gaps 10;

Qy	97	GGCCAGCGCGGGCTGTGCGCGCTCGCACACCGGAACGAGCTGTGTCGATCCGGG	156
Db	43	GGTCACCGTGGCTCCCGCGGGCTCGACATCGCGGACGAGCTGTCATGTCGCGC	102
Qy	157	CCACGAGACCCCAAGCC---GTCGCGGGCGCTGAGCGGGCAGGTGTGCAGATATCGG	213
Db	103	CCGCGAGGACCGAACCAGGTCCCGAGGCGCGACGTGAAGACGTGCGCGTGTGCGC	162
Qy	214	CGACGAGTGGGTCTACCGTGGACGCGACCTTCTGTCGCCCTGCAACGAGTGGCGCTT	273
Db	163	CGACGAGTGGGACGCGGGAGGACGCGGAGCCCTTCTGTCGCGTGTGCGCGGCTT	222
Qy	274	CCCGTGTGCGGCCCTGCTACGAGTACGAGCGCGGAGGCGACGAGTATCGG	333
Db	223	CCCGCTGTGCGGCCCTGCTACGAGTACGAGCGCGGAGGCGACGAGTGTGCGCGCA	282
Qy	334	GTGCAAGACGCGCTACAAGCGCCCTCAAGGGAGCGCGAGGTTGCCGGGACGATACGA	393
Db	283	GTGCAACACCGCTACAAGCGCCGAGAAAGGTGCGCGAGGTTGAAAGGGACGAGGAGGA	342
Qy	394	GG---AGGACATCGACACCTGGAGCACGAGTTCACATCGACGACGAGATCAGCAGAG	450
Db	343	GGGCGCGGAGATGACGACTTCGAGGACGAGTTCGCCCGCAAGAGCCCTCA	402
Qy	451	GCAGCTGGAGGCGAATGACGAGACGACGATCAGCGGCGATGTGTCAGCGCAGGAT	510
Db	403	CGAGCTGTGCGCGTTCGAGCTTACTCGAGAACGCGGCGACACCGGCGCAGAAATGGG	462
Qy	511	GAGCTACGCGAGGGGCGCCGACGACGCGGCAACAAACACCGCGAGATCCGCGCCAT	570
Db	463	GACGGGTG-----CAGAGCTGTGCTC	486
Qy	571	CATCACCGCTCCGCTCCGTCCGTGAGCGGTGAGCGGTGAGTTTCCGATTACCAACGGGTATGG	630
Db	487	CTTACCGGAAGCGTCGCCGGGAAG-----	512
Qy	631	CCAGGCGAGGTCTCGTCTTCCCTGCAAGCGCATCCATCCGTACCTGTGTGTAGCC	690
Db	513	-----ACCTGGAGCGCGAGAG	528
Qy	691	AGGAGTGCACAGTGGGACGAGAAGAAGTGAAGTGAAGGAGGAGGATGACAGACTG	750
Db	529	GGAGATGGAGGGGAGCATGAGTGAAGGACCGGATCGACAGTGAAGACCAAGACGGA	588
Qy	751	GAAAGTCAAGCAGGCGATCTTCGCGCGCGCGCGATCCCGAAGACATGAGCAGCGCAGCT	810
Db	589	GAAAGGGGCAAGCTCAACACGACGACGCGACGACGACGACGACGACGACGACGACGACG	648
Qy	811	G---GCACTGAACGACGAGGGAGGACGCGCTGTGAGGAAGAGTGTGATCGCGTCGAG	867
Db	649	GTACATGCTGTTGCCGAGGCGGACAGCGCGCTGTGCGCAAGGTTCCGATCCCGTCGAG	708
Qy	868	CAAGGTGAACCGTACCGGATGTGATCGTGTGCGTCTGCTGTGCTCGCTCTTCTCCT	927
Db	709	CATGATCAACCGTACCGCATGTCATCGTGTGCTCGCGCTGCTGCTGTGCTGTGCTCTTCTCCT	768
Qy	928	CCGCTACCGTATCCTGCAACCCCGTCCCGGACGCGCATCGGGCTGTGGCTCGTCTCCATCAT	987
Db	769	CAAGTTCGCGATCAGCAGCGCGCCACGAGCGCGCTGTGTGCTGCGCTCGCTCAT	828
Qy	988	CTCGGAGATCTGGTTCGCCATCTCTCTGGATCTCTGACACGATTCGCCAAGTGGTTCGCCAT	1047

Db	829	CTCGAGCTCTGGTTCGCCCTTCTCTGGATCTCTGGACAGCTGCCAAAGTGGGCGCCGCT	888
Qy	1048	CGACCGCGAGACGTACCTCGACCGCCCTCTCCCTCAGGTACGAGAGGAAAGGAGCGGTC	1107
Db	889	GACCGGGAGACGTACCTGGACCGCTGCGGTACGACCGTGAAGGCGGCGTG	948
Qy	1108	GCTGTGTGCGCGGTGGACCTGTGTGAGACAGTGTGAGACCGCTCAAGGAGCGCGCT	1167
Db	949	CCGGCTGTCTCCCATCGACTTCTTGTGACAGGTGGACCGCTCAAGGAGCGCGCAT	1008
Qy	1168	GGTGACCGCAACACCGTGTCTTCTTCTGCGGTAGATACCCCTGAGCAAGGTCTC	1227
Db	1009	CATACCGCCAAACCGTGTCTTCTTCTGCGGTGACTTACCCCTGAGACCGCGTCAG	1068
Qy	1228	CTGTACGTCTCGAGCGCGCTCGATGTCTGACGTTTCTGAGTGTGTCGAGAGAGGCG	1287
Db	1069	CTGTACGTCTCGAGCGCGCGCTTCTTCTGACGCGCTTCTGACGCGCTGTCGAGACGC	1128
Qy	1288	CGAGTTCGCGCGCAAGTGGGTGCGCTTCTGCAAGAGTTCGGCATCGAGCCCGCGGCC	1347
Db	1129	CGAGTTCGCGCGCGCTGGGTGCGCTTCTGCAAGAGTTCGCGGTGAGCGCGCGGCC	1188
Qy	1348	GGAGTTCCTACTTCTCGCTCAAGTCTGACTACTCTCAAGGCAAGGTGCGAGCCCTTCGT	1407
Db	1189	GGAGTTCCTACTTCTGCGAAGATCGACTACCTCAAGGCAAGGTGCGAGCGCGCTTCGT	1248
Qy	1408	GCAGGACGCGCGCCATGAAAGAGAGTATGAGGAGTTCAGAGTTCAGAGTTCACACGCGCT	1467
Db	1249	CAAGGACGCGCGCCATGAAAGAGGAGTACGAGGAGTTCAGGTGCGCATCAACGCGCT	1308
Qy	1468	GTTGGCCAAAGCCATGAAGTTCGCGGAGAGGTGATCATGAAGACGCGACGCGCTG	1527
Db	1309	GGTGGCCAAAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1368
Qy	1528	GCCCGGGAACAACACCGCGACACCCCGCATGATCCAGGTTCCTGCGGCGACAGCGG	1587
Db	1369	GCCCGGGAACAACACCGCGACACCCCGCATGATCCAGGTTCCTGCGGCGACAGCGG	1428
Qy	1588	CGGCCACGACCGAGGGCAACGAGTTCGCCCGCTCGTGTGATCTCTCCGTGAGAGAGG	1647
Db	1429	CGCGTGGAGTGGAGGGCCACAGCTGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1488
Qy	1648	CCCGGATTCAGACACCAAGAGGCGCGCGCATGAACGCTGTGATTCGCGTCTCGC	1707
Db	1489	CCCGGGTACAACACCAAGAGGCGCGCGCATGAACGCTGTGCGCGCTCTCTCTCTCTCTCT	1548
Qy	1708	CGTGTGACCAACGCGCCATTCATGCTCAACTTGGATGTGATCTACATACACACAG	1767
Db	1549	CGTGTGACCAACGCGCCCTTCATCTCAACCTCGACTGCGACCATCTACGTCAACACAG	1608
Qy	1768	CAAGGCGCATCCGGGAGGCGCATGTGCTTCTCTCATGGACCTTCAGTTCGCGCGGAGGTCTG	1827
Db	1609	CAAGGCGGTGCGGAGGCGCATGTGCTTCTCTCATGGACCTTCAGTTCGCGCGGAGGTCTG	1668
Qy	1828	CTAGTTCAGTTCGCGGAGGTTGACGCGCATTCGACGTCGACGACGACGACGACGACGACG	1887
Db	1669	CTAGTTCAGTTCGCGGAGGTTGACGCGCATTCGATGCGATCGATCGATCGATCGATCGAT	1728
Qy	1888	GACACCGTCTTCTTCTGACATCAACATGAAGGGGTGAGCGCATCCAAAGGCGCGGTGA	1947
Db	1729	CAACGTCGTCTTCTTCTGACATCAACATGAAGGGGTGAGCGCATCCAAAGGCGCGGTGA	1788
Qy	1948	CGTGGGACAGGGGTGCTGTTCCGCGCGCGAGGCGCTCTACGGGTACAAACCTTCCCAAGG	2007
Db	1789	CGTGGGACAGGGGTGCTGTTCAACCGCGAGGCGGTGTACGGGTACGACCGCGCGCGGCC	1848
Qy	2008	ACCAAGAGCGCCAAAGATGTGAC-----TGCGACTGTGTGCGCGCT	2049
Db	1849	CGAGAGCGCGCCAAAGATGACGTGCGACTGCTGCGCGCTGCTGTGCTGTGCTGTGCTGTG	1908
Qy	2050	CTTTCGCGCGCAAGAGG-----	2067

Db 1909 CTTGCGCGGGCAAGCGCGGCAAGCGCGCAAGGACAAGAGGGCGACGCGGGCGAGGA 1968
QY 2068 -----GAAACACGCCAAGAGCGG 2085
Db 1969 GCGCGCGCGGGGCTGCTCGGCTTCTACAGGAAGCGGACAGAGGACAAGCTCGGCGG 2028
QY 2086 GCTGCCGAGGGGACCGCTGATATATGGAGTAGATAGCGACAAGAG-----2131
Db 2029 CGGGTCGGTGGCGGCGAGCAAGAGGGCGCGGGCTGTACAGAAGACACAGCGCGCGTT 2088
QY 2132 -----ATGCT 2136
Db 2089 CGAGCTGGAGGAGATCGAGAGGGGCTGGAGGGGTACGAGCGCTGGAGCGCTCTCGCT 2148
QY 2137 CATGTCCCAATGAATTCGAGAAGCGGTTTCGGGCAGTCCGGCGGCTTCGTTCACGCTCGAC 2196
Db 2149 CATGTCGACAGAAGAGCTTCGAGAAGCGGTTTCGGCCAGTCGCCGTGTTTCATCGCCTCCAC 2208
QY 2197 GCTGATGGAGGAAGGGCGGCTCCTC-----CTTGGTCGAGCCCGCGCGCTCCTCAA 2250
Db 2209 GCTCGTCGAGGACGGCGGCTGCGCAGGGCGCGCGCGGACCCCGCGCGCTCATCAA 2268
QY 2251 GGAGGCCATCCATGTTCATCAGCTGCGGCTACGAGGACAGACCGACTGGGGCTGGAGCT 2310
Db 2269 GGAGGCCATCCAGCTTCATCAGCTGCGGATACGAGGAGAGACCGAGTGGGGCAAGAGAT 2328
QY 2311 GGGGTGGATCTAGGGTGATCAGCGAGGACATCTCTGACGGGGTTCAAGATGCACTGCGG 2370
Db 2329 TGGGTGGATCTATGGGTGCTGACAGAGGATATCTCTGACGGGGTTCAAGATGCACTGCGG 2388
QY 2371 CGGGTGGCGCTCGGTACTGATGCGGAAGGGGGGGGTTCAAGGGTTCGGCGCGGAT 2430
Db 2389 GGGGTGGAAATCGGTACTGACACGCGACACGGCGGGGTTCAAGGGTTCGGCGCGCAT 2448
QY 2431 CAATCTATCGGACCGTCTCAACAGGTGCTCGGTGGCGCTGGGGTTCGGAGATCTT 2490
Db 2449 CAATCTGTCTGATGCTTCCACAGGTGCTCGGTGGCGCTGGGGTTCGGAGATCTT 2508
QY 2491 CTTTCAGCGCGCACAGCCCCCTGTGTATCGGCTACAGAGACGGCAACTCAAGTGGCTGA 2550
Db 2509 CATGAGCCGCACTGCCGCTCGGTACGCTAC-----GGCGCGCGCTCAAGTGGCTGA 2565
QY 2551 GGGCTTGGCTATACAAACACACATCTACCCCTTTCACCTGCTCCCGCTGCTGCTGCTA 2610
Db 2566 GGGCTTGGCTATACAAACACACATCTGATGATACCCCTTTCACCTCCATCCCGCTCTCGCTA 2625
QY 2611 CTGACCCCTCCCGCGCTGCTGCTCTACCGCAAGTTTCATCATGCGCTGATAGCAC 2670
Db 2626 CTGCACATCCCGCGCTGCTGCTCTACCGCAAGTTTCATCATTTCCACGCTGACAA 2685
QY 2671 GTTTCGACGCTTCTTTCATCGCCCTCTTTCATGTCCATCTTTCGCGACGGGATCCTGA 2730
Db 2686 CTTTCGACGATCTGGTTTCATCGCGCTCTTCTGTTTCATCATCGACGAGCTCCTGA 2745
QY 2731 GATCGGTTGAGCGGGGTGAGCATCGAGAGTGGTGGAGAACGACGAGTTCTGGGTAT 2790
Db 2746 GCTGCGGTGGAGCGGGGTGAGCATCGAGGACTGGTGGCGCAACGACGAGTTCTGGGTCT 2805
QY 2791 CGCGCGGCTGCGCGCATCTTTCGCGCTTCGTCGAGGGCTGCTCAAGGTCTCGCGG 2850
Db 2806 CGCGCGGCTGTCGCGCATCTTTCGCGGTGTTTCAGGGCTTCTCAAGGTTCTGGCGG 2865
QY 2851 GATCGACACCAATTCACCGTCACTTCAAGGCCACCGCGACGAGGACGACGAGTTTCG 2910
Db 2866 CGTGGACACAGCTTCACCGTCACTTCAAGCGCGCGGCGAGGCGCGCTTCG 2925
QY 2911 CGAGCTCTAGGCTTCAAGTGGACACGCTCTCATCTCCGCGCCACACGCTGCTCATAT 2970
Db 2926 GGACCTCTACCTTTCAGTGGACCACTGCTGCTGGTGGCGCCCGCCACACGCTCATCAT 2985
QY 2971 TAACTGTCATCGGCGTGGCGGATCTCCGACGCCATCAACAAAGGGTACAGTCCG 3030
Db 2986 CAACATGGTGGGATCTGGCGCGGCTGTCCGACGCCGCTCAACACCGGCTACGGCTCCTG 3045

QY 3031 GGGGCCCTCTTCGGCAAGCTCTTCTTCGCTTCTGGGTATCGTCCACCTCTACCGTT 3090
Db 3046 GGGCGCGCTTTCGGCAAGCTCTTCTTCTCTTCTGGGTATCGTCCACCTCTACCGTT 3105
QY 3091 CTTCAAGGGGCTCATGGGGCGCAAGACAGGACGCCACCGTGTGTGTATCATCTGTGTCAT 3150
Db 3106 CTTCAAGGGGCTCATGGGGAGGAGAACCGGACGCCACCATCGTGTGTCTGTGTCAT 3165
QY 3151 TCTGTGGGCTCATCTTCTCTCTCTCTGGGTAGGATCGACCTTTCATCGTCAGGAC 3210
Db 3166 CTTCTCTCGCTCATCTTCTCTCTCTGGGTAGGATCGACCTTTCATCGTCAGGAC 3225
QY 3211 CAAGGGCCGCGAGCTGAGCGAGTGTGGCATCAATTCCTGAGCT 3253
Db 3226 CAAGGGCCCATCTCAAGCCATCGGAGTCGAGTCTGAGCT 3268

BT013780 3414 bp mRNA linear PLN 11-MAY-2004
Lycopersicon esculentum clone 132670R, mRNA sequence.
DEFINITION
BT013780
ACCESSION
BT013780.1 GI:47105195
VERSION
FLI_CDNA.
SOURCE
Lycopersicon esculentum (tomato)
ORGANISM
Lycopersicon esculentum
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 3414)
REFERENCE
Kirkness, E.F., Wang, W. and Vazeille, A.
AUTHORS
Direct Submission
TITLE
Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
JOURNAL
Medical Center Drive, Rockville, MD 20850, USA
FEATURES
Location/Qualifiers
1..3414
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="132670R"
/tissue_type="fruit"
/note="TZXY89"

ORIGIN
Query Match 45.1%; Score 1552.6; DB 8; Length 3414;
Best Local Similarity 69.7%; Pred. No. 7.3e-195;
Matches 2209; Conservative 0; Mismatches 909; Indels 51; Gaps 6;
QY 89 GCGATGGAGGCGAGCGCGGCTGTGGCGGCTCGCAACCGGAAACGAGTGTGTG 148
Db 63 GCGATGGAGGCTAGTGCAGGCTAGTGTGCTGCTTACATAACCGGAATGAACCTGTGCTC 122
QY 149 ATCCGGGCGCACAGAGCCCAAGCGCTGCGGGCGCTGAGCGGCGAGTGTGCGAGATA 208
Db 123 ATTCATGGCCATGAAGAGCATAAAGCCATTGAAGGATTTGAGTGGACAAGTTGTGACATA 182
QY 209 TCGCGCGACAGGCTCGGGCTCACGCTGGACGCGGACCTCTTCGTCGCTCAACGAGTGC 268
Db 183 TGTGGGATGAATAGGCTTAACAGTGGATGATCTTTTGTGGCTTGCATGAGTGT 242
QY 269 GGCCTTCCCGTGTGCGGCCCTTGCTACGAGTACGAGCGCCGGAGGGGCAACGAGAACTGC 328
Db 243 GGTTCCTCTGTGTGCGGCTTGTACGAGTATGAGAGAAGGAAGGAAGAACTCAACAGTGT 302
QY 329 CCCAGTGCAGACCGCTTACAGCGCTCAAGGGGAGCCCGAGGGTTCGCCGGGACGAT 388
Db 303 CCACAGTGCAGACCGAGATACAAAAGTCTTAAAGGAGTCCAAAGGGTGGCGGAGATGAT 362
QY 389 GACGAGGAGGACATCGACGACCTGGAGCACGAGTTCAACATCGACGACGAGAATCAGCAG 448
Db 363 GATGAGGAGATATTGACGATATTGACATGATTTCAAGTTGATGATGA-----412

QY 449 AGGCAGCTGGGGCAACATGCACAAACAGCCAGATCAACGAGGCGATGCTGCACGGCAGG 508
DB 413 -----CCAAAAAAGAAATAGAAACATTTGTTGAGACCAATCTCCCATGTTAAG 458
QY 509 ATGAGCTACGGGAGGGCCCGACGACGCGGCAACACACCCCGCAGATCCCGCCC 568
DB 459 ATGACATATGGAGAGACCAAGATG-----AAGATTCCGCTCAGTACCCCTCCT 509
QY 569 ATCATCACGGCTCCCGTCCGTTCCGTTGAGCGGTGAGTTTCGATTAACAAC---GGG 625
DB 510 GTCAATTCTGCAACCCGCTACATCCGTTAGCGGTGAATTCGCCAATATCGAACCATGGA 569
QY 626 TATGGCCACCGGAGGTCTCGTCTTCCCTGCACAGCGCATCCATCCGTACCTGTGTCT 685
DB 570 AATGGAGACAGATGTTGGGATCTTCGCTTCAAAAAGGATACATCCCTATCCAGCTTCT 629
QY 686 GAGCCAGGAGTGCAAAGTGGGACGAGAAAGAGTGAAGTGAAGGAGAGAGATGGAC 745
DB 630 GAATCTGGAAGTGAAGATGGATGATAGAAAGAGGGGTTGGAAGAAAGNATGAA 689
QY 746 GACTGGAAGTCCAAGCAGGGGATCCTCGGCGGCGCCGATCCGGAAGACATGGACGCC 805
DB 690 GATTGGAAAATTCAGCAAG---ACACGCGGGGCAAGATTATGATGACTCTGCTGACGTT 746
QY 806 GACGTGSCACTGAACGAGGAGGCGAGCGCGCTGTGAGGAAGGTGTCGATCGCGTGC 865
DB 747 GATATGTCCATGGTGGATGAAGCAAGACAGCCTTTATCAAGAAAGTGCCTAATGATCA 806
QY 866 AGCAAGGTGAACCCGTACCGATGGTGTATCGTGTGCTCTCGTTGCTCGCCTCTTC 925
DB 807 AGCAGATCAATCTTACCGATGGTGTATGTAGTAGCTAGGTGGTGTGATTCGCGCTCTTC 866
QY 926 CTCGGTACCGTATCTCTGCAACCCCGTCCCGAAGCCATCGGGCTGTGGCTCGTCTCCATC 985
DB 867 CTTGCTTATAGAATTTCTGAATCCAGTGCATGATGCGATGGTCTCTGTTAACTTCTATC 926
QY 986 ATCTGCGAGATGTTGGCATCTCTCGATCTCTCGACCATGTTCCCAAGTGGTTCCCC 1045
DB 927 ATCTGTGAGATCTGGTTTGGCTTCTCTTGGATCTCTGATCAGTTTCCCAATGGTTCCCT 986
QY 1046 ATCGACCGCAGAGCTACCTCGACCGCTCTCCCTCAGGTACGAGAGGAGGGAGCGG 1105
DB 987 ATCGACCGTGAAGCTTACCTTGTATGCTCTCTCAGGTATGAGAGGAGGAGAACCA 1046
QY 1106 TCGTGTCTGTGCGGGTGGACCTGTTCTGTGAGCAACGTTGACCCGCTCAAGAGAGCCCGG 1165
DB 1047 AATATGCTAGCCCGGTAGATGCTTTGTCAGTACAGTGGATCCTATGAAGGAACCTCT 1106
QY 1166 CTGGTGACCGCAACACCGTCTCTCCATCTCGCCGTAGACTACCCCGTGGACAGGTC 1225
DB 1107 CTTGTTACAGCTAATACAAATCCTCTCAATATTAGCAATGGATTACCCCTGTTGACAAAT 1166
QY 1226 TCCTGTACGTCTCCGACGAGCGGCGTGCATGCTGACGTTTCGAGTCTGCTCGGAGAGC 1285
DB 1167 TCATGTTACCTATCTGATGTGAGCTTCAATGTGCACCTTTGAGCACATATCAAAACT 1226
QY 1286 GCCAGTTCGCGCAAGTGGTGGCTCTCTGCAAGAAAGTTTCGGCATCGAGCCCGCGCC 1345
DB 1227 GCAGAAATTTGTAGAAAATGGTTCCTCTGCAAGAAAGTTTGTATAGAGCCCTCGAGCT 1286
QY 1346 CCGAGTTCATCTTCTCGCTCAAGTGCATCACTCAAGCAAGGATGAGGTCAGCCCACTTC 1405
DB 1287 CCAGAAATTTTACTTCTCCCTAAAGATTGATTTATCTCAAGGACAAAATCCAGCCAACTTC 1346
QY 1406 GTGACGAGCGCGCGCATGAGAGAGATGATGAGGAGTTCAGGTCCCGATCAACGCG 1465
DB 1347 GTTAAGGAGAGACGAGCAATGAAGAGAGATGACGAGAAATTCAGGTAAAGAAATCAATGCA 1406
QY 1466 CTGTGTCGAAGGCATGAAGGTCCCGCAGAGGGGTGGATCATGAAGGACGCGACGCGG 1525
DB 1407 CTAGTTGCCAAAGCAACTAAATGCTCTCCGAGGATGGATCATGCAAGATGGAACACCA 1466
QY 1526 TGGCCCGGGAAACAACCCCGGACCAACCCCGGCATGATCCAGGTGTTCTCTGGGCGCAGC 1585

DB 1467 TGCCAGGAATAATACAGGGATCATCTCTGGTATGATTCAAGTCTTTTTAGGCCAAAGT 1526
QY 1586 GGGCGGCACACACCGAGGGCAACGAGCTGCCCGCTCGTGTACGTCTCCCGTGAAG 1645
DB 1527 GAGGAACTGATGTGGATGACATGACTCCCTGCTTGTATGTTTCTCGTGAANA 1586
QY 1646 CGCCCGGATTCAGACACCAAGAGGCGCGGCCAATGAACGCTCTGATTCGCGTCTCC 1705
DB 1587 AGGCTGGTTTCCAGCATCAAAAAGGCTGGTGTATGAACGCCCTGGTTCGTTGTTCA 1646
QY 1706 GCGTGTGACCAACCGCCATTCATGCTCAACTTGGACTGTGATCACTACATCAACAAC 1765
DB 1647 GAGATTCTTACCAATGACCAATTTATGCTTAACTTGGATTGTGATCAATACCTTAAACAAC 1706
QY 1766 AGCAAGGCCATCCGGAGGCGCATGTCTCTCATGAGACCTCAGGTTCGCGCGGAAGGTC 1825
DB 1707 AGCAAGGCTCCAGAGAGGCCATGTCTTCTTGTATGGACACACAAATGGGCAAAAAGTT 1766
QY 1826 TGCTAGCTTCAGTTCCGCGAGAGGTTGACGCGGATCGAGTCGACGACCGATACGCTAAC 1885
DB 1767 TGCTTCTGTCAGTTTCTCAAAGATTTGATGGAATAGAACGACGATAGATATGCCAAC 1826
QY 1886 AGGAACACGCTCTCTTCGACATCAACATGAAGGGCTGACCGCATCCAGGCCCGGTG 1945
DB 1827 CGAAACACAGTCTCTTTTGATATTAACATGAAAGGCTTAGATGGACTTACAAGGTCAGTA 1886
QY 1946 TAGCTGGGACAGGCTGGTGTTCGCGCGCACGGCTCTACGCTTACAAACCTCCCAAG 2005
DB 1887 TATGTTGGCACAGATGTGTTTCAGNAGCAGCATTATATGGCTACATCTCTCTTAA 1946
QY 2006 GGAACCAAGAGGCCAAGATGTTGACCTGCGACTGTGCGCTGCTTCGCGCCGCAAGAG 2065
DB 1947 CGTGCCAAAGCGCCAAAGATGGTGGCTGCGACTGTGTCATGTTTGGACGCAAAAG 2006
QY 2066 CGG-----AAACACGCCAAGAGCGGCTGCCCGAGGGCACCGCTGATATGGGNGTAGAT 2119
DB 2007 AAGCTTTGATAAATATAAATCTGAAGTCAATGCAAGATGACGACAAATGCTAAGGATTTGAC 2066
QY 2120 AGGCAAGAGAGATGCTCATGTCCCATCGAACTTCGAGAGCGGTTTCGGGCGAGTCCGCG 2179
DB 2067 GATGACNATGAGTACTATATGTCAGATGACATTTGAAAAGAAATTTGGTCACTGTCGA 2126
QY 2180 GCGTTCGTCACTGCGCTGATGGAGAAAGCGGCTCCCTCTCTCGTTCGAGCCCGCC 2239
DB 2127 ATTTTGTGACATCAACTTTAATGATTGAAGTGGGTTCCACTTCTGTCGAGCCACGA 2186
QY 2240 GCGCTCTCAAGAGGCCATCCATGTATCATGCTGCGGCTACGAGGACAGACCGACTGG 2299
DB 2187 GCCTTGTGAAGGAAGCTATCCATGTGATCAGCTGTGGCTATGAAGATAAAACAGAAATGG 2246
QY 2300 GGGCTGGAGCTGGGTGGATCTACGGGTGATCAACGAGGACATCTCTGACGGGTTCAAG 2359
DB 2247 GGTGTTGGAGCTGGATGGAATACGGTTCTATCAGAGGATATCTTAAAGGTTTCAAG 2306
QY 2360 ATGCACTGCGCGGGTGGGCTCCGTGTACTGATGCGGAAGCGGGCGGCTTCAAGGGG 2419
DB 2307 ATGCACTGCGGTGGAGTCACTTATGATGCTTAAATTTAGCTGATTAAGGGG 2366
QY 2420 TCGCGCCGATCAATCTATCGGACGCTCTCAACAGGTGCTCCGGTGGGCTGGGTC 2479
DB 2367 TCTGCTCCAATCAATCTATCAGATCGTCTCAACAGGTGCTTCGATGGGCTCTTGGTTCG 2426
QY 2480 GTGAGATCTTCTTACGCGGCAACGCCCCCTCTGTGTACGGCTTACGAAGACGGCAACCTC 2539
DB 2427 GTTGAGATCTTTTCTAGTCAATCATAGCCAGTTTGGTATGGCTATAGGAGGAAATCTT 2486
QY 2540 AAGTGGCTGAGCGCTTCGCTTACATCAACACCAACCATCTTACCCCTTCACTCGCTCCCG 2599
DB 2487 AAGTGGCTTGAACGATATCATACATTAACACAACTATTTTACCCCTTCACTTCTTTACCT 2546
QY 2600 CTGCTGCTTACTGACCTTCCCGCGCTCTGCTCTCTACCGGCAAGTTTCATCATGCGG 2659

Db	2547	CTCCTTGCCCTACTGCACTCTCCCTGCTGCTGTTGCTCACTGCGAAAAATTCATAATGCCA	2606
QY	2660	TGCATTAGCAGGTTCCGCCAGCCTCTTCTTCATCGCCCTCTTCATGTCCTCATCTTCGGCAGC	2719
Db	2607	GAGATAAGCAGCTAGCAAGTCGTGTTTTCATGCTCTTTCTCTCTATCTTACACGC	2666
QY	2720	GGCATCTCGAGATCGGTGGAGCGGGTGAGCATCGAGAGTGTGGAGGAACGACGAG	2779
Db	2667	GGAAATCTTGGAGTAAAGTGGAGTGAGTAAGCATCGAGAAATGGTGGAGAAATGAACA	2726
QY	2780	TTCTGGGTATCGGCGCGGTGTCGGCGCATCTCTTCGCGGTGTCGAGGCGCTGCTCAAG	2839
Db	2727	TTTGGGTATGGAGGTGTATCTGCCACCTTTTTCGTGTAGTACAAGGTCTTCTCAAG	2786
QY	2840	GTCTCGCGGGATCGACACCACTTTCACGCTCACCTCAAGGCCACCGCGCAGGAGAC	2899
Db	2787	ATTTGGCTGGCATAGTACCACTTATCTGTGACATCTAAGGCAACATGATGAGGAC	2846
QY	2900	GACGAGTTCCGCGAGCTTACGCTTCAAGTGGACACACCTCTCTATCCCGCCACCG	2959
Db	2847	-----TTCGAGAGTTATATGCTTCAAGTGAGCTACTCTCTTATTCCTCAACA	2900
QY	2960	CTGCTCATATTAACCTCATCGGCTGTCGGCGGCATCTCCGACGCCATCAACACGG	3019
Db	2901	ATTCTCATATCAACTTGGTGGGTGTAGCCGGAATCTCTGACCTATCAACACGGA	2960
QY	3020	TACCAGTCTCGGGGCCCTCTTCGGCAAGCTCTTCTCGCTTCTGGGTCTATCGTCCAC	3079
Db	2961	TACATTATGGGTCCCTTTCGGAAGCTTTCCTTGGCTTTGGTTCATGTTCTAT	3020
QY	3080	CTCTACCCCTTCTCAAGGGCTCATGGGCGGCCAGAACAGGACGCCACCGTGTGTC	3139
Db	3021	TTATATCCATCTCTCAAGGCTCTATGGCGAGCAGAACAGGACTCCAACCATTTGTGA	3080
QY	3140	ATCTGCTCATCTGCTGCTCCATCTTCTCCCTCTCTGGTCAAGTCCACCTTTC	3199
Db	3081	ATATGCTCATCTTTTATGCTCCATCTTTCTGCTCTGGGTTCGAATTGATCCATTT	3140
QY	3200	ATCGTCAGGACCAAGGGCGCGACGTCAGGCACTGTGGCATCAATTGCT	3248
Db	3141	GTGCTCAAAACCAAGCACCAGATGTCAAGCATGTGGAGTGAACGTGCT	3189
RESULT 8			
AY139754			
LOCUS	AY139754 3355 bp mRNA linear PLN 02-SEP-2002		
DEFINITION	Arabidopsis thaliana AT5g17420/T10B6_80 mRNA, complete cds.		
ACCESSION	AY139754.1		
VERSION	AY139754.1 GI:22654964		
KEYWORDS	FLI CDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids 1; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 3355)		
AUTHORS	Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y., Huan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A. and Ecker, J.R.		
TITLE	Arabidopsis cDNA clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 3355)		
AUTHORS	Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y., Huan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A. and Ecker, J.R.		

TITLE	JOURNAL		
COMMENT	RIKEN Genomic Sciences Center (GSC) Consortium members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.		
FEATURES	source		
1..3355	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/chromosome="5"		
	/clone="RAFL09-35-F05 (R22199)"		
	/ecotype="Columbia"		
46..3126	/note="cellulose synthase catalytic subunit (IRX3)"		
	/codon_start=1		
	/product="AT5g17420/T10B6_80"		
	/protein_id="AA098075.1"		
	/db_xref="GI:22654965"		
	/translation="MEASGLVAGSHNRNELVVIHNEBPKLNKLDGFCBICGDOI GLTVEGDLFVACNEGFPACPCYEVEEREGTQPCQCKTVKRLGRSFRVSGDEDEE DIDDIEYFNIEHQDKKHSABMLYKMSYGRPEDDENGFRFPVIAGHSGEFPV GGGYNGEHLGHRVHPYPSSSAGSEGGRRMRDDWKQHGNGLGEDDDPENGLIDE ARQPLSRKVPKPIERETVLDRLSLRYEREGERENMLAPVDFVSTVDPKPEPLVT FAVSWILAMDYPVEKISCYVSDGASMLTPESLSEAFARKWPKFCCKFSEIRAP EMYFTLVQDLYQVHPTFVKERRAMKREYEEFKVIRNAQVAKSKVPLEGIMODGT FMPGNNTKDHGMIQVFLHSGGDFVEGHELPRLVYVREKRPQGHKKKAGAMNALV RVAGVLTNAPFMLNLDCHYVNNKAVREAMCFMDPQIGRKVCYVQFPQRFEDGDTN DRYANRNTVFFDINNKGDLGGVYVGTGCVFKRQALYGYEPKPKRPMISGCC PCFGRNRNKKPSKNDMNDVAALGAGSGDKHLMSEWNFEKTFQSSIFVSTLMEE GGVPPSSPAVLLKEAIIHVISGYEDKTEWGLGWIYGSITEDILTGPKMKCRGWS TYCMPKPAFKGSAPINLSDRNLQVLRWALGSVEIFFSRHSPLWYGYKGGKWLKWERF AYANTITPFTSIPLLAYCILPAICLLDKFIMPPISTIPASLFFSILFMSIIVTGILE LRWSGVSTSEWNRNEQFVIGGISAHLFVQGLLKLIGADITNFVTYSKATDDDDFG ELYAGFTWTLTLPPTVLIINIVGVAGISDAINNGYSWGLFGLKLPFSFWIVHLY PFLKMGQRNRTPTIIVTWSVLLASIFSLMWLRIDPFLVKTKGPDTSKGCINC"		
Query Match	44.1%;	Score 1519.2;	DB 8; Length 3355;
Best Local Similarity	69.8%;	Pred. No. 1.8e-190;	
Matches 2213;	Conservative 0;	Mismatches 863;	Indels 96; Gaps 8;
QY	88	CGCGATGAGGCGCAGCGCGGCTGTGTGCGCGGCTCGCACACCGAACGAGCTGTGCT	147
Db	42	CCCTATGGAAGTAGCGCGGCTGTGTGCGCGGCTCTATAACCGTATGAAGTGTGCT	101
QY	148	GATCGGGGCCACGAGGACCCCAAGCCGCTGCGGGCGCTGAGCGGCAGGTGTGCGAGAT	207
Db	102	CATTACAAACCATGAAGACCAAGCCCTCTGAAGAATCTAGATGGACAAATCTCTGAGAT	161
QY	208	ATGCGCGCAGAGGTGCGGCTCACCGTGGACCGGACCTCTTCGTCCTCGACACGATG	267

Db 162 ATGTGGAGATCAGATTGGTTTAA CAGTAGAGGAGACCTCTTCGTAGCTTGCATAGATG 221
Qy 268 CGGCTTCCCGTGTGCCGCGCTCTGCTACGAGTACGAGCGCGGAGGCGACGCAAACTG 327
Db 222 TGGTTTCCGCGGTGTAGACCTTGCTATAGTACGAGAGAGAGAAACACAAAACG 281
Qy 328 CCCCAGTGAAGACGCGCTACAGCGCTCAAGCGCTCAAGGGGAGCCGAGGTTGCCGGGACGA 387
Db 282 TCCTCAGTGAAGACTCGTTACAAGCGCTCAGAGGAAGCCCAAGAGTGGAGGAGATGA 341
Qy 388 TGACGAGGAGACATCGACGACCTTGGAGCAGCAAGTTCAACATCGACGACGAGATCAGCA 447
Db 342 AGACGAGAGATATTGATGATATGAGTATGATTTAATATCGAATGACACAGA--- 398
Qy 448 GAGGACGTGAGGGCAACATGACAGACAGCCAGATCACCGAGCGGATGTCGACGCGAG 507
Db 399 -----TAAGCATAGCAATTCGTGAGGCTATGCTTTATGGAA 437
Qy 508 GATGAGCTACGGAGGGGCCCGAGCAGCGGCGACCGCAACACACCCCGCAGATCCCGCC 567
Db 438 AATGAGCTATGGAAGAGGTCTGAGGATG-----ATGAGAAATGGGAGATTCACACC 488
Qy 568 CATCATCACCGGCTCCGCTCCGCTGAGCGGTGAGTTCCCGATTACCAACGGGTA 627
Db 489 TGTATAGCTGGTGTCA-----TAGTGGAGAAATTCAGATTGGAGGAGTTA 536
Qy 628 TGGCCACGGGAGGTCTCGTCTTCCCTGCAACGCGCATCCATCCGTACCCCTGTGCTGA 687
Db 537 TGGTAATGGAGAAC-----ATGGGCTTCATAAGCGGTGCAACCCATATCCATCTGA 590
Qy 688 GCCAGGAGTGCCTAAGTGGGACGAGAGAAAGATGAGCTGGAAAGAGAGGATGGACGA 747
Db 591 AGCTGGGAGT-----GAGGGAGGATGGCGGGAAGAATGGATGA 629
Qy 748 CTGGAAGTCCAAAGCAGGCGATCTCGCGCGCGCGCATCCCGAAGACATGGACGCGA 807
Db 630 CTGGAAGTCCAGATGGAATCTTGGCCGACGACGATGATGATC-----CTGA 680
Qy 808 CGTGGCACTGAACGACGAGCGAGCGAGCGCTCTCGAGGAAGGTGCGATCGCGTCGAG 867
Db 681 GATGGGACTGATCGACGAGGACGCGACGCACTCTCGCGGAAAGTTCCTATTCGCTCAAG 740
Qy 868 CAAGGTGAACCCGTACCGATGGTGTGATCGTGGTGGTCTCGTTGTGCTCGCTTCCTCT 927
Db 741 CAAGATCAATCCATATCGGATGGTGCATCGTGGTGGTCTGATGATCTAGCAGTTTCT 800
Qy 928 CCGGTACCGTATCTGCAACCCGCTCCGCGAGCGCATCGGGCTGTGGCTCGTCTCCATCAT 987
Db 801 CGGATATAGGCTCTTGAATCAAGTGCATGATGCTCTGGGATATGGCTGACCTCTGTGAT 860
Qy 988 CTGCGAGATCTGGTTCGCCATCTCTCGACATCTCGACAGTTCGCCCAAGTGGTTCCCAT 1047
Db 861 CTGTGAATCTGGTTCGCTGCTCTTGGATCTTGTATCAGTTCCTCCCAAGTGGTTCCCTAT 920
Qy 1048 CGACCGGAGACGTACCTCGACCGCTCTCCCTCAGGTA CGAGAGGAAGGGAGCGCTC 1107
Db 921 TGAACGTGAGACCTATCTAGATCGGCTTTCCTCAGGTACGAGAGAAAGGTGAACCAA 980
Qy 1108 GCTGCTCTGCGCGGTGACCTGTTGTGAGCAGCGTGAACCGCTCAAGGAGCCCGCGCT 1167
Db 981 TATGCTTGGCCCTGTAGATGCTTTGTGACGATCGGTGGACCCATTGAAGGAGCCTCCCT 1040
Qy 1168 GGTGACCGCAACACCGTGTCTCCATCTCGCGGTAGACTACCCCGTGGACAGAGTCTC 1227
Db 1041 CGTCACATCCAAACACTGTGCTGTCATCTTGGCCATGGAATACCCAGTTGAGAAATCTC 1100
Qy 1228 CTGCTAGCTCTCGACAGCGCGGTGATGCTGAGCGTTGAGTTCGCTGCGAGAGCGG 1287
Db 1101 CTGCTATGTCTCTGACGACGCGTGTCAATGTCTTACATTCGAATCTCTCTCGGAAACTGC 1160
Qy 1288 CGAGTTCGCGCAAGTGGGTGCCCTCTTCGACGAAGTTTGGCATCGAGCCCGCGGCC 1347
Db 1161 TGAGTTTGCAGAAATATGGGTTCCTCTCTGTGAAGAAATTCCTCATAGAGCCCGGCC 1220

Qy 1348 GAGTTCCTACTTCTCGCTCAAGGTGCACTACCTCAAGGAACAAGTGCAGGCCCACTTCGT 1407
Db 1221 GGAATGTACTTACGTTGAAAGTTGATTATCTCAGGACAAAGTCCACCCCAATTTGT 1280
Qy 1408 GCAGGCGCGCGCCATGAAGAGAGATGATGAGAGTTCAAAGTTCGGGATCAACGGCT 1467
Db 1281 TAAGGAACGTGAGCCATGAAGAGATATGAGAGTTCAGGTGAGATCAACGCTCA 1340
Qy 1468 GGTGGCCAAAGCCATGAAGTTCGCGCAGAGGGGTGGATCATGAAGGACGCGACGCCGTG 1527
Db 1341 AGTGGCGAAGCCCTCAAGGTTCTCTAGAAGGTTGGATCATGCAAGATGGAACACCGTG 1400
Qy 1528 GCCCGGAAACAACACCCCGGACCAACCCCGGATGATCAAGGTTCCTTGGGCGCACAGGG 1587
Db 1401 GCCAGGGAACAACCAAGGACCAACCCCGGTATGATCCAAGTCTTCTCGGCCACAGCGG 1460
Qy 1588 CGGCCACGACACGAGGGGCAACGAGCTGCCCGCTCGTGTACGCTCTCCGTGAGAGCG 1647
Db 1461 AGGATTTGATGTGGAAGGCGATGAGCTTCTCGGCTTGTGTGTGTGCTCCGTGAGAGCG 1520
Qy 1648 CCGGGATTCAGCACCACCAAGAGGCGCGCCATGAACGCTCTGATTCGCGCTCTCCG 1707
Db 1521 TCTGTGTTTCAACACCAAGAAAGCTGGGCGCATGATGCCCTGCTTTCGAGTGGCAG 1580
Qy 1708 CGTGTGACCAACGCGCCATTCATGTCTCAACTTGGACTGTGATCACTACATCAACAACAG 1767
Db 1581 CGTACTCACAATGCTCTCTTTCATGCTGAACCTTGGACTGTGATCACTATGTAAACAACAG 1640
Qy 1768 CAAGGCCATCGGGAGGCGATGCTTCTCTCATGAGCCCTCAGGTCCGCGCGGAGGCTG 1827
Db 1641 CAAGGCCGTGAGGGAAGCAATGTGTTTTTGTGATGATCCTCAGATTTGAAAGAGGCTG 1700
Qy 1828 CTAGTTCAGTTCGCGCAGAGGTTCCAGCGCATCGAGCTGCACACGATACGCTAACAG 1887
Db 1701 CTATGTTCACTTCCTCAAGGTTGATGCAATTGACACAACGATGTTACGCGCAACAG 1760
Qy 1888 GAACACCGTCTTCTTCAACATCAACATGAAGGGCTGGAGCGCATCAAGGCCCGGTGTA 1947
Db 1761 AAACACAGTCTTCTTGTATCAATATGAAGGTCTAGATGGAATCAAGGTCCAGTTTA 1820
Qy 1948 CGTGGGACAGGTGCGTGTTCGCGCCAGGCGCTCTAGCGCTACAAACCTCCCAAGGG 2007
Db 1821 CGTTGTACTGTGTGTTTTTCAACGACAGCTCTGTATGTTTATGAACCAACCAAGGG 1880
Qy 2008 ACCCAAGAGCCCAAGATGCTGACTGCTGCCGCTGCTTCGCGCGCAGAGAGCG 2067
Db 1881 TCCTAAAGTCCAAAGATGATAGCTGTGTTGTTGCTTCTTGGCGCCGAGGAA 1940
Qy 2068 GAAACA-----CGCCAAAGACGGGCTGCGGAGGGCACCGCTGATATGGAGTAGA 2118
Db 1941 GAATAAGAAATTTTCCAAGAATGACATGAATGGTAGCTAGCAGCCCTTGGAGGAGCAGA 2000
Qy 2119 TAGGACAAAGAGATGCTCATGTGCCACATGAACCTTCAGAGAGCGGTTCCGCGAGTCCG 2178
Db 2001 AGGTGATAAGGAGCATTTGATGTTGTTGTAATGAACTTTGAGAAAACATTTGGGCAATCATC 2060
Qy 2179 GGGCTTCGTCAAGTGCAGCTGATGGAGGAGGGCGGCTCCCTCTCTCTCGTGCAGCCCGC 2238
Db 2061 CATCTTTGTAACCTCAACTTTGATGGAGAAAGGTGGTGTCTCTCCCTCATCAAGTCTGTC 2120
Qy 2239 CGGCTCTCTCAAGGAGCCATTCATGTTCATCAGTCTGGGTTCAGAGGACAGACCGACTG 2298
Db 2121 AGTGTCTCTTAAAGAGGCAATCCATGTCATAAGCTGCGGTTATGAAGACAAGACTGAGTG 2180
Qy 2299 GGGGCTGGAGCTGGGTGGATCTACGGGTGATCAGGAGGACATCTCTGACGGGTTCAA 2358
Db 2181 GGGAACTGAGCTGGGTGGATCTATGGCTCTATCACAGAGGATATTTGACGGGATTCAA 2240
Qy 2359 GATGCACTGCGCGGGTGGCGCTCCGCTGTACTGATGCCAAGCGCGCGGCTTCAAGGG 2418
Db 2241 GATGCACTGCGGTGGATGGAGGCTATTTTACTGATGCTTAAAGGCGCTGCATTCAAAGG 2300

QY	2419	GTGGCGCCGATCAATCTATCGGACCGTCTCAACACGAGTGTCTCGGTGGGCGCTGGGGTC	2478
Db	2301	TTGAGTCTCTTAATTAATCTATCAGACAGAGTTAAACAGGTTTTCGGTTGGGCACTTGGATC	2360
QY	2479	CGTCGAGATCTTCTTTCAGCGCGGACAGCCCGCTGCTCTAGCGCTACAGAAACGCGACCT	2538
Db	2361	GGTTGAGATATTTTTTCAGCGCGGACAGTCTCTCTGTTATGGCTACAAAGAGGAGCAAACT	2420
QY	2539	CAAGTGGCTGGAGCGCTTCGCTCATCAACACACCAACCATCTACCCCTTCACCTCGCTCCC	2598
Db	2421	CAAGTGGCTGGAGCGCTTTCGCTATGCGAACACCAACCAATCTACCCCTTCACATCTATACC	2480
QY	2599	GCTGCTCGCTCTACGACACCTCCCGCGCGTCTGCTCCTCACCGGCAAGTTCAATCATGCC	2658
Db	2481	ACTTCTTGGCTTACTGTATCTTTCAGCGCATCTGTCTCTACTGACAAAATTCATCATGCC	2540
QY	2659	GTGATATTAGACAGTTCGCGAGCCTCTTCTATCATCGCCCTCTTCAATGCTCAATCTTCGGAC	2718
Db	2541	ACCGATTAAGACATTTCTGCTAGTCTCTCTTCTCATCTCACTGTTATGTCGATCATTTGAAC	2600
QY	2719	GGGCAATCTGGAGATGCGGTGGAGCGGGGTGAGCATCGAGAGTGGTGGAGGAACAGACA	2778
Db	2601	GGGAATCTTGGAAATTGAGATGGAGCGGAGTTAGCATTTAGGAAATGGTGGAGAAACAGACA	2660
QY	2779	GTTCTGGGTCAATCGCGCGGTGTCGCGGCAATCTCTTCGCGGTGTCGAGGGCTGTCTCAA	2838
Db	2661	ATTCTGGGTCAATTTGGAGGAATCTCAGCTCATCTCTTTGCGGTTGTCTCAAGGTCTCTCTCAA	2720
QY	2839	GGTCTCTCGCGGGATCAGACCAACATTTCACGGTCACTCTCCAGGCCACCGCGACGAGGA	2898
Db	2721	AATCTTAGCAGGCATTGACAAACTTCAACCGTCAATCAAGGCAAC-----AGATGA	2774
QY	2899	CGACGAGTTCGCGGAGTCTACGCGCTTCAAGTGGACCAACGCTCTCTCATCCGCCACACAC	2958
Db	2775	TGATGACTTTGGAGAACTTTACGCAATCAATAGACAAACACTGCTGATCCTCTCAACACAC	2834
QY	2959	GCTGCTCATATTAACTCATCGCGCTGTCGCGGCACTCTCCGACGCCATCAACACACG	3018
Db	2835	TGCTTAATCATAAACATTTGTTGGCGTTGTTGACGGCATCTCAGATGCCATTAAACAATGG	2894
QY	3019	GTACCACTCTGGGGGCGCTCTTCGCGAGCACTCTTCTCGCTTCGCGTCTGAGTCACTGCTCA	3078
Db	2895	ATATCAGTCTTGGGAGCTCTATTGTTGTAACCTCTCTCTCTCTTTTGGGTCAATGTTCA	2954
QY	3079	CCTCTACCGCTTCTCTCAAGGGGCTCATGGGGCGCAGAACAGGACGCCACCGTGTGTTGT	3138
Db	2955	TCCTACCCATCTCTCAAGGTCTGATGGGTAGACAGAACAGAACCAACCAATTTGGT	3014
QY	3139	CATCTGTCTCATCTCTGTCGCGCTCCATCTTCTCCCTCTCTGCGTCAAGGATCGACCCCTT	3198
Db	3015	GATTGGTCTAGTGTATTGCGCATCTATCTTCTCTTTGTTGGGTAAGAATTGATCCTTT	3074
QY	3199	CATGCTCAGACCAAGGCGCGGACGTCAGGCAGTGGGATCAATTTGCTGA	3250
Db	3075	TGTGCTCAAGACCAAGGACCTGACACTTCCAAGTGTGGCATCAACTGCTGA	3126
RESULT	9		
BT004543			
LOCUS	BT004543	3081 bp	linear
DEFINITION	Arabidopsis thaliana AT5g17420/t10B6_80	gene, complete cds.	PLN 19-FEB-2003
ACCESSION	BT004543		
VERSION	BT004543.1	GI:28416516	
KEYWORDS	FLI CDNA		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE			
AUTHORS			

Db 2020 TTTGTAACTTCAACTTTTGATGGAAGAGTGGTGTTCCTCCGTCATCAAGTCTCCAGTG 2079
Qy 2243 CTCTCAAGAGGCCATCATGTGTCATGCTGGCTACGAGCAAGACCGACTGGGG 2302
Db 2080 CTCCTTAAGAGCAATCATGTGTCATGCTGGCTTATGAGCAAGACTGAGTGGGA 2139
Qy 2303 CTGAGCTGGGTGGATCTACGGGTGATCACGGAGGACATCTCTGACGGGGTTCAAGATG 2362
Db 2140 ACTGAGCTGGTTGGATCTATGCTCTATCACAGAGATATTTTGACGGGATTTCAAGATG 2199
Qy 2363 CACTGCGGGGTGGCGCTCGTGTACTGATGCCAAGCGGCGGCTTCAAGGGGTG 2422
Db 2200 CATTCGCGGTGGATGAGGTCTATTACTGATGCCCTAAGAGGCGCTCAATTCAAGGTTCA 2259
Qy 2423 GCGCGCATCAATCTATCGACCGTCTCAACGAGTGTCTCGGTGGCGTGGGGTCCGTC 2482
Db 2260 GCTCCTATTATCTATCAGACAGTTAAACGAGTTTGGGTGGGACTTGGATCGGT 2319
Qy 2483 GAGATCTTCTCAGCGGCAAGCCCGCTCTCTGCTAGCGGTACAAAGAACCGCAACCTCAAG 2542
Db 2320 GAGATATTTTACGCGGACAGTCTCTCTGATGCTTACAAAGAGGCAACTCAAG 2379
Qy 2543 TGGCTGAGCGCTTCGCTTACATCAACACACCATCTACCCCTTCACTCGTCCCGCTG 2602
Db 2380 TGGCTTACGCGTTTGTCTATGCAACACAAATCTACCCCTTCACTATACCACTT 2439
Qy 2603 CTGCGCTACTGACCTCCCGCGTCTGCTCTCCTCAGCGGAGTTCATCGCGCTG 2662
Db 2440 CTGCGCTACTGATCTCTCCGCGCATCTGCTCTTACTGACAAATTCATATGCCACCG 2499
Qy 2663 ATTAGCAGCTTCGCCAGCTCTTCTTCTATCGCCCTCTTCTATGTCATCTTCGCGAGCGG 2722
Db 2500 ATAAGCACATTTGCTAGTCTCTTCTTCTATCTCACTGTTTATGTCGATCATTTAAACGG 2559
Qy 2723 ATCTGAGATGCGGTGGAGCGGGGTAGCATCGAGAGTGTGGAGGAACGAGCAGTTC 2782
Db 2560 ATCTTGAATTGAGATGGAGCGGAGTTAGCATTTAGAGAAATGGTGGAGAAACGAGCAATTC 2619
Qy 2783 TGGGTCTATCGCGCGGTGTCGCGCATCTCTTCCCGCTGCTGAGGCGCTGCTCAAGTTC 2842
Db 2620 TGGGTCTATGGAGAAATCTAGCTCATCTCTTGGCGGTGTCCAGGTCTCTTCAAAATC 2679
Qy 2843 CTCGCGGATGACACCACTTACCGTCACTTCAAGCGCACCGCGACGAGGACGAC 2902
Db 2680 TTAGCAGGCATTGACACAACTTACCGTCACTCAATCAAGGCAAC-----AGATGATGAT 2733
Qy 2903 GAGTTCGCGAGCTCTACGCTTCAAGTGACACCGTCTCTATCCGCGCCACACGCTG 2962
Db 2734 GACTTTGGAGAACTTTACGCAATTTCAATGGACAAACACTGCTGATCCCTCCCAACACTGTC 2793
Qy 2963 CTCATCATTAACGTCATCGCGCTGCGCGGCGCATCTCCGACGCCATCAACACGGGTAC 3022
Db 2794 TTAATCATAAACATTTGTCGCGTGTGTCAGGCACTCTCAGATGCCATTAACAATGATAT 2853
Qy 3023 CAGTCTCGGGGCGCCCTCTTCGCAAGCTCTTCTTCCGCTCTTGGGTCTATCGTCCACCTC 3082
Db 2854 CAGTCTTGGGACCTCTATTGTTAACTCTTCTTCTTCTTGGGTCTATGTTCTATCTC 2913
Qy 3083 TACCGTTCCTAAGGGGCTCATGGGCGCGCAGACGAGACGCCACCGTGTGTTGTCATC 3142
Db 2914 TACCCATTCCTCAAGAGTCTGATGGGTAGACAGAACAGAACACCAACCATTTGTTGAT 2973
Qy 3143 TGTGTCATTTGCTGCTCCTCATCTCTCTCTCTCTCTGCTGCTGAGTCCAGCCTTTCATC 3202
Db 2974 TGTGAGTGTATTGGCATCTATCTCTCTTCTTCTTGGTGAAGAAATGATCTCTTTGTG 3033
Qy 3203 GTCAGGACCAAGGGCGCGAGCTGAGGATGAGTCAATTTGCTGA 3250
Db 3034 CTCAGACCAAGGACCTGACACTTCCAAGTGTGGCATCAACTGCTGA 3081

RESULT 10

AF088917 Arabidopsis thaliana 3081 bp mRNA linear PLN 02-JUL-1999
LOCUS Arabidopsis thaliana cellulose synthase catalytic subunit (IRX3)
DEFINITION mRNA, complete cds.
ACCESSION AF088917
VERSION AF088917.1 GI:4886755
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1 (bases 1 to 3081)
AUTHORS Taylor,N.G., Scheible,W.R., Cutler,S., Somerville,C.R. and
Turner,S.R.
TITLE The irregular xylem3 locus of Arabidopsis encodes a cellulose
synthase required for secondary cell wall synthesis
JOURNAL Plant Cell 11 (5), 769-780 (1999)
MEDLINE 99264300
PUBMED 10330464
REFERENCE 2 (bases 1 to 3081)
AUTHORS Taylor,N.G., Poindexter,P., Scheible,W., Cutler,S., Somerville,C.R.
and Turner,S.R.
TITLE Direct Submission
JOURNAL Submitted (31-AUG-1998) Biological Sciences, 3.614 Stopford
Building, University Of Manchester, Oxford Road, Manchester M13
9PT, UK
FEATURES
source Location/Qualifiers
1..3081 /organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="Landsberg erecta"
/db_xref="taxon:3702"
/chromosome="V"
1..3081 /gene="IRX3"
/note="cellulose synthase catalytic subunit"
1..3081 /gene="IRX3"
/codon_start=1
/product="cellulose synthase catalytic subunit"
/protein_id="AAD32031.1"
/translation="MEASAGLVAGSHRNELVVIHNEEPKPLKNDGGCFCEICGDOI
GUTVGDLFVACNEGQFPACPCYERREGTQNCQCKTRYKRLGSRVGEDEDE
DIDDEYBNIEHQDKHSAEAMLYKMSYGRGDEDDENGFRFPVIAGHSGEPV
GGYNGEHLKRVHPVPSSEAGSGGWRERDDWKLOHGNLGPEDDDPEMGLIDE
AQPLSRKPIASSKINPYRMVIVARLVAVFLRYRLNPNVHDALGLMTSVICEM
FAVSWILDQFPKFPPIERTYLDRLSLRYEREGEPNMLAPVDVFTVDPKPEPLV
SNVLSILAMDYPVEKISCYSDDGASMLTFESLSEAFARKWPFCKFSIEPRAP
EMVFTLKVDLQDKVHPTFKERRAMKREYEFKVRINAQVAKASKVPLEGWTMQDGT
PWPNTKDHGMIQVFLHSGGFDVEGHELPLVYVSRKRPFGQHKKAGAMNALV
RYAGVLTAFMNLCDHYVNSKAVREAMCFMADPOIGKVCYVQFPORPDITN
DYARNNTVFFDINMKLDGIQGPVTGTCVFKQALQIYEPKPKPKPKMISCGCC
PCGRRRNKFKSKNDMNGDVAALGGAGDEKHELMFEMNFEKTFGSSIFVSTLMBE
GGVPSPPSAPVLLKEAIVHISGDEKTEWGLWYIGSTEDILTITGFMHGRS
IVCMRPAKPAKAPINLSDRLNVLWALGSVEIFPSRSLPWLKYGKGLKWLRF
AYANTVPTSTIPLLAYCILPAICLLTDKTFMPPISFTASLFFISLEWSIIVTGILE
LRWSGVSEWRNRPQFWIGISAHLEFANVQGLLKLILAGIDTNFTVTSKATDDDFG
EYIAFKWITLLIPPTTLLIINIVVAGISDALNNGYQSWGLFGKLFPSFWVHY
PFLGLMGRQNETPTTIWVWSVLLASIFSLMLWRVIDPFVLTKGPDTSKGINC"

Query Match 44.1%; Score 1516.8; DB 8; Length 3081;
Best Local Similarity 69.8%; Pred. No. 3.8e-190;
Matches 2210; Conservative 0; Mismatches 862; Indels 96; Gaps 8;
Qy 92 ATGAGGCCAGCGCCCGGCTGTGGCGGCTCGCACACCGAACGAGCTGGTCTGATC 151
Db 1 ATGGAAGCTAGCGCGGCTCTGTGCGCGGTTCTCATAACCGTAATGAACCTAGTCGTCATT 60
Qy 152 CGGGGCCACGAGGACCCCAAGCCGCTGAGCGGCGGAGGTGTGCGGAGATGCG 211

Db 61 CACAACCATGATGAGCCAAAGCCTCTGAAGAACTCTAGATGACAAATTCCTGTGATATGT 120
Qy 212 GCGACAGAGGTGCGGCTCAGCGTGGACGGCGACCTCTTCGTCGCTGCAAGAGTGGCGC 271
Db 121 GGAGATCAGATTTGGTTTAAACAGTAGAAGGAGACCTCTTCGTAGCTTGCATAGATGGT 180
Qy 272 TTCCCGGTGTGCGGCTCTCAGAGTACGAGCGCGGAGGAGCGACGCAAGTGGCCCC 331
Db 181 TTTCCGGGTGTAGACCTTGTATGATACGAGAAAGAGAAAGAAACACAAAATGTCTCT 240
Qy 332 CAGTGAAGACGCGCTCAAGCGCCTCAAGGGGAGCCGAGGGTTGCGGGGACGATGAC 391
Db 241 CAGTGTAAAGCTCGTTACAGGCTCTCAGAGAAAGCCCAAGAGTGGAGGAGATGAAGAC 300
Qy 392 GAGAGGACATCGACGACCTGGAGCAAGAGTTCAACATCGACGACGAGAATCAGCAGAG 451
Db 301 GAAGAAGATATTGATGATATTGATGATGAATTTAATATCGAAATGAACAAAGAAAA 353
Qy 452 CAGCTGAGGGCAACATGCGAAGCAGCCAGATCAACGAGGCGATGCTGCACGGCAGGATG 511
Db 354 -----TAAGCATAAAGCATTTCTGCTGAGGCTATGCTTTATGGGAAATG 396
Qy 512 AGCTACGGGAGGGCCCGACGACGCGCGACGCGCAACAAACACCCCGCAGATCCCGCCCATC 571
Db 397 AGCTATGGAAGAGTCTCTGAGATG-----ATGAAATGGGAGATTTCCACCTGTT 447
Qy 572 ATCACCGGCTCCCGCTCCGTCGCGGTGAGCGGTGAGTTCCGATTAACACGGGTATGGC 631
Db 448 ATAGCTGGTGTCA-----TAGTGGAGAAATTTCCAGTTGGAGGAGTTATGGT 495
Qy 632 CAGGCGAGGTCTCGTCTTCCCTGCACAGGGCATCCATCCGTACCCCTGTGTCTGAGCCA 691
Db 496 AATGGAGAAC-----ATGGGCTTCATAAGCGGTGTGCACCCCATATCCCATCTCTGAAGCT 549
Qy 692 GGGAGTCCCAAGTGGGACGAGMAAGAAAGTGAAGTGGAGGAGAGATGGAGACTGG 751
Db 550 GGGAGT-----GAGGAGGATGGCGGAAAGATGGATGACTGG 588
Qy 752 AAGTCCAAAGCAGGGCATCTCGGCGGCGCGCGATCCGGAAGACATGGACGCGACGTG 811
Db 589 AAGCTCAGCATGGAATCTTGGCCAGAACACAGATGATGATC-----CTGAGATG 639
Qy 812 GCATGAAACGACGAGGCGAGCAGCCGCTGTGAGGAAAGTGTGATGCGGTGAGGCAAG 871
Db 640 GGAATGATGACGAGGACGCGACGACCACTCTCGCGAAAGTTCCCATTTGCTCAAGCAAG 699
Qy 872 GTGAACCCGTACCGGATGGTGAATCGTGTGCTCTGTTGTGCTGCTCTTCTCTCCGG 931
Db 700 ATCAATCCATATCGGATGGTCAATGCTGTAGGCTTGTGATTTAGCAGATTTTCTCGGG 759
Qy 932 TACCGTATCTGCAACCCCGTCCCGAGCCCATCGGGCTGTGGCTCGTCTCCATCATCTGC 991
Db 760 TATAGGCTTTGAATCCAGTGATGATGCTCTGGGATTTATGGCTGACCTCTGTGATCTGT 819
Qy 992 GAGATCTGGTTGGCCATCTCTGATGATCTGTGACAGTTCCCAAGTGGTTCCCATCGAC 1051
Db 820 GAAATCTGGTTGGTGTCTCTGGATTTCTGATCAGTTTCCCAAGTGGTTCCCTATTGAA 879
Qy 1052 CGCGAGACGTACCTCGACCGCTCTCCTCAGTACGAGGAGGAGGAGCGGAGCGCTCGCTG 1111
Db 880 CGTGAGACCTATCTAGATCGGCTTCTCCTCAGGTACGAGAGAGAGTGAACCAATATG 939
Qy 1112 CTGTCGCGGTGGACCTCTGTTGTGAGCACGAGTGGACCCGCTCAAGGAGCGCGCTGGTG 1171
Db 940 CTGGCCCTGTGATGCTTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 999
Qy 1172 ACGCCCAACACCGTGTCTCATCTCTCGCGTGAAGTACACCGTGGAGCAAGGTCTCTGTC 1231
Db 1000 ACATCCAAACATGTGTCTCAATCTTGGCCATGACTACCCAGTTGAGAAAAATCTCTGTC 1059
Qy 1232 TAGTCTCCGACGAGCGGCTGATGCTGAGCTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 1291
Db 1060 TATGTCTCTGACGAGCGGTGCTTCAATGCTTACATTTCAATTTCTCTCTCGGAAACTGCTGAG 1119

Qy 1292 TTCCGCGCGCAAGTGGGTGCGCTTCTGCAAGAAAGTTTCGGCATCGAGCCCCCGCGCCCGGAG 1351
Db 1120 TTTGCAAGAAAAATGGGTTCCTTCTGTAAGAAATTTCTCATAGAGCCACCGGCGCACCGGAG 1179
Qy 1352 TTCTTACTTCTCGCTCAAGGTTCGACTACTCAAGGACAAAGGTGACAGCCACCTTCGTGCGAG 1411
Db 1180 ATGTACTTTCACGTTGAAAGTTGATTATCTCCAGGACAAAGTCCACCAACATTTGTTAAG 1239
Qy 1412 GAGCGCGCGCCCATGAAGAGAGAGATGAGGAGTTCAAGGTTCGGATCAACCGCTGTGTG 1471
Db 1240 GAAGCTCGAGCCATGAAGAGAGATATGAGGAGTTCAAGGTTCAGATCAACGCTCAAGTG 1299
Qy 1472 GCAAGGCCATGAAGAGTTCGCGCAGAGAGGTGAGATCATGAAGGACGCGCATCGCTGGCCC 1531
Db 1300 GCGAAGCCCTCAAAGGTTCTCTAGAAAGTTGGATCATGCAAGATGGAACAACCGTGGCCA 1359
Qy 1532 GGGAAACAACCCCGACACCCCGGCATGATCCAGGTGTTCTTGGGCCACAGCGGGCGC 1591
Db 1360 GGGAAACAACACCAAGGACCCCGGTATGATCCAAAGTCTTCTCGGCGCACAGCGGAGGA 1419
Qy 1592 CAGACACACGAGGGCAACGAGCTGCCCGCTCTGTTACGTCTCCCGTGAGAACGCGCGC 1651
Db 1420 TTTGATGTCGAAGGGCATGAGCTTCTCGGCTTGTGTACGTGTCCTCGTGAGAAAGGTCTCT 1479
Qy 1652 GGAATTCAGCAACCAAGAAAGCGCGGCATGAACCGTCTGTGATTCGGGTCTCCCGCGTG 1711
Db 1480 GGTTTTCAACACCAACAAAGAGTGGCGCATGAATGCCCTGTTTCGAGTGGCAGCGTA 1539
Qy 1712 CTGACCAACCGGCATTCATGCTCAACTTGGACTGTGATCACTACATCAACAACAGCAAG 1771
Db 1540 CTCACAAATGCTCTCTTTCATGCTGAACCTTGGACTGTGATCACTATTAACAACAGCAAG 1599
Qy 1772 GCCATCCGAGGCGCATGTGCTTCTCATGAGACCCCTCAGGTTCGCGCGGAGGCTCTGTAC 1831
Db 1600 GCCGTGAGGCAAGCAATGTGTTTTTGTGATGATCTCTCAGATTCGAAAGAGGTCTGCTAT 1659
Qy 1832 GTTCAGTTCCCGCAGAGGTTTCGACGGCATCGACGTGCAACGACCATACGTAACGTAACAGGAC 1891
Db 1660 GTTCAGTTCCCTCAAAGGTTTGTGTCATTTGACAAACAGCATCTGTTACGCCAACAGAAAC 1719
Qy 1892 ACCGTCTTCTTCGACATCAACATCAAGGGCTGAGCGGCATCCAAAGCCCGGTGTGAGTC 1951
Db 1720 ACAGTCTTCTTGTATATCAATATGAAGGTCTAGATGAAATCCAAAGTTCAGGTTTACGTT 1779
Qy 1952 GGGACAGGTCGCTGTTCCGCGCGCAGCGCTCTACCGCTACAAACCTCCCAAGGAGACCC 2011
Db 1780 GGTACTGTTGTGTTTTCAACGACAGCTCTGTATGTTATGAACCAACCAAGGTCCT 1839
Qy 2012 AAGAGGCCCAAGATGGTGAACCTGCACTGTGCGCGGTGCTTGGCGCGCAAGAGCGGAAA 2071
Db 1840 AAACGTCCAAAGATGATAAGCTGTGGTTGTTGCTTGTCTTGGCGCGCGGAGAAAGAA 1899
Qy 2072 CA-----CGCCAGGAGCGGCTCCCGAGGGCACCCCTGATATGGGATGATAGC 2122
Db 1900 AAGAAATTTTCCAAAGATGACATGAATGGTGAAGTAGAGCCCTTGAGGAGCAGAAAGT 1959
Qy 2123 GACAAGGAGATGCTCATGTCCCATGAACTTCGAGAAAGCGGTTTCGGCGAGTCGCGCGCG 2182
Db 1960 GATPAAGGAGCATTTGATGTTTGAATGAACTTTGAGAAACATTTGGCAATCATCCATC 2019
Qy 2183 TTGCTCACTGCACTGATGGAGGAAAGCGGCGTCTCTCTGTCGAGCCCCCGCGCG 2242
Db 2020 TTTGTAACCTCAACTTTGATGAGAAAGTGGTGTCTCTCGTCATCAAGTCTCTGCACTG 2079
Qy 2243 CTCTCAAGAGGCCATCCATGTGTCAGCTCGGCTACGAGGACAGACCGATGGGG 2302
Db 2080 CTCTTAAAGAGGCAATCCCATGTATGAAGTGGGTTTATGAAGCAAGACTGAGTGGGA 2139
Qy 2303 CTGAGCTGGGTGGATCTACGGGTGATCAACGAGGACATCTCTGAAGGGGTTCAAGATG 2362
Db 2140 ACTGAGCTGGGTGGATCTATGGCTCTATACAGAGGATATTTTGAAGGATTCAGATG 2199

Qy	2363	C	A	T	G	C	C	G	G	G	T	G	G	C	T	C	G	T	G	T	A	C	T	G	C	A	G	C	G	G	C	G	T	T	C	A	A	G	G	G	T	C	A		2422																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
Db	2200	C	A	T	T	G	C	G	T	G	A	T	G	A	G	G	T	C	A	T	T	A	C	T	G	C	A	T	A	A	G	G	C	T	G	C	A	T	T	C	A	A	G	G	T	C	A		2259																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										
Qy	2423	G	C	G	C	G	A	T	C	A	T	C	G	G	A	C	C	G	T	C	A	A	C	A	G	S	T	C	G	G	T	G	G	G	C	G	T	G	G	G	T	C	G	T	C		2482																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
Db	2260	G	C	T	C	T	A	T	A	T	C	T	A	G	A	C	A	G	T	T	A	A	C	C	A	G	S	T	T	T	G	G	T	T	G	G	T	T	G	G	G	C	A	T	T	G	A	T	C		2319																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																								
Qy	2483	G	A	G	A	T	C	T	T	C	A	G	C	G	G	C	A	G	C	C	C	C	T	G	T	A	C	G	S	T	C	A	A	G	A	C	A	G	A	C	A	C	T	C	A	A		2542																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
Db	2320	G	A	G	A	T	T	T	T	C	A	G	C	C	G	C	A	G	A	C	A	G	T	C	T	C	T	G	T	A	T	G	G	C	T	A	A	G	A	G	A	C	A	A	C		2379																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
Qy	2543	T	G	G	C	T	G	A	G	C	G	T	T	C	G	C	T	A	C	A	A	C	A	C	A	C	A	C	A	C	A	C	A	C	T	A	C	C	C	T	C	A	C		2602																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
Db	2380	T	G	G	C	T	G	A	G	C	G	T	T	T	G	C	T	A	T	G	C	C	A	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C		2439																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
Qy	2603	C	T	C	G	C	T	A	T	G	C	A	C	C	T	C	C	C	C	G	T	C	T	C	A	C	A	C	G	G	A	A	G	T	T	C	A	T	C	A	T	C		2662																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
Db	2440	C	T	T	G	C	T	A	C	T	T	C	C	A	G	C	C	A	C	T	G	T	C	C	T	A	C	T	G	A	C	A	A	T	T	C	A	C	A	A	T	T	C	A		2499																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
Qy	2663	A	T	T	A	G	A	C	A	G	T	T	C	C	T	C	T	C	A	T	C	C	T	C	A	T	G	C	C	T	C	A	T	G	T	C	A	T	T	C	G	A		2722																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
Db	2500	A	T	A	A	G	A	C	A	T	T	G	T	C	T	C	T	C	T	C	T	C	T	C	A	T	G	T	T	A	T	G	T	C	A	T	T	A	T	G		2559																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
Qy	2723	A	T	C	T	G	A	G	A	T	C	G	C	G	G	T	C	A	G	A	T	C	G	A	G	A	T	C	G	A	G	A	T	C	G	A	G	A	T	C	G	A		2782																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
Db	2560	A	T	C	T	T	G	A	A	T	T	G	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	A		2619																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
Qy	2783	T	G	G	T	C	A	T	C	G	G	C	G	T	T	C	G	C	A	T	C	T	T	C	G	C	G	T	C	G	A	G	G	C	T	G	C	T	C	A	A	G	T		2842																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
Db	2620	T	G	G	T	C	A	T	T	G	A	G	A	A	T	C	T	C	A	G	T	C	A	T	C	T	T	T	G	C	G	T	T	G	C	A	A	G	T	C	C	T	C		2679																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
Qy	2843	C	T	C	G	C	G	G	A	T	C	G	A	C	C	A	C	T	C	A	C	C	T	C	A	A	G	C	C	A	C	C	G	C	A	G	A	G	A	C		2902																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
Db	2680	T	T	A	C	A	G	C	A	T	T	G	A	C	A	A	A	C	T	T	C	C	G	T	C	A	C	A	T	C	A	A	A	G	G	C	A	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

RESULT 11	AK072259	3426 bp	mRNA	linear	PLN 24-JUL-2003
LOCUS	AK072259				
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:J023009D02, full insert sequence.				
ACCESSION	AK072259				
VERSION	AK072259.1	GI:32982282			
KEYWORDS	FLI_CDNA; CAP trapper.				
SOURCE	Oryza sativa (japonica cultivar-group)				
ORGANISM	Oryza sativa (japonica cultivar-group)				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Magnoliopsida; Liliopsida; Poales; Poaceae; SpERMatoPhyta; Magnoliophyta; Liliopsida; Poales; Poaceae;				

Koya,S., Kurihara,C., Mateuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Satoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.
Location/Qualifiers
1. .3426
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023009D02"

FEATURES

source

ORIGIN

Query Match	43.4%	Score 1494.6	DB 8	Length 3426
Best Local Similarity	70.3%	Pred. No. 3.1e-187		
Matches 2296	Conservative 0	Mismatches 669	Indels 300	Gaps 10
QY	187	GAGCGGCGAGTGTGCGAGATATCGCGCGACGAGTGGGCTCACGTTGACGCGACCT 246		
Db	72	GAAGGAGAGACGTGCCGGTGTGCGGCGAGAGTGGCGGAGGAGCGGAAGCC 131		
QY	247	CTTTCGTCCTCAACAGTTCGGCTTTCCCGTGTGCGGCGCTGTCTACGAGTACGAGCG 306		
Db	132	GTTTCGTGGCGTGGCCGAGTTCGGCTTTCCCGTGTGCAAGCCCTGTACGAGTACGAGCG 191		
QY	307	CCGGGAGGCGACGAGAACTGCCCCGAGTGCAGACGCGCTACAAGCGCTCAAGGGGAG 366		
Db	192	CAGCGAGGCGACCCAGTGTCTGCCCCGAGTGCACCCCGTACAAGCGCCCAAAAGGGTG 251		
QY	367	CCCAGAGGTTGCCGGGACGATACGA---GGAGGACATCGACACCTGGAGCACGAGTT 423		
Db	252	CCCACGGTGGAGCGACGAGGACGCGCGCGCATGACGACCTTCGAGGAGAGATT 311		
QY	424	CAACATCGACGA CGAGAAATGACGAGAGCGAGCTGGAGGGCAACATGCGAGAACGCGAGAT 483		
Db	312	CCAGATCAAGAGCGCCCAACGACGA-----AACCCGCCACGAGCCCGTC 357		
QY	484	CACCGAGCGATGTGCAAGCGACGATGAGTACGGGAGGGGCCCCGACGACGCGCAGCG 543		
Db	358	AACCTTCAGCTTACTTCGAGAACGGCGAGACGCGGCAAGAAGTGGCGCCCTTGAGAG- 416		
QY	544	CAACAACACCCCGCAGATCCCGCCCATCATCACCGGCTCCCGCTCCGTCGCGGTGAGCGG 603		
Db	417	-----CCCGCGCTCTCTTCTTTCACCGGAAGCGTGGCTGGGAGG- 457		
QY	604	TGAGTTCCGATTACCAACGGGTATGCCACGGCGAGGTCCTGCTTTCCTTCGCAACAGCG 663		
Db	458	----- 457		
QY	664	CATCCATCCGTACCCTGTGTGAGCCAGGAGTGCCAAATGGGACGAGAGAAAGAGT 723		
Db	458	-----ATCTGGAGCAGGAGGAGATCGAGGTTGGCATGAGTGGAGGACAGATCGA 512		
QY	724	GAGCTGAAGAGAGGATGAGACGATCGGAGTCCAGCAGGGGCATCTCTCGCGCGCGCGCG 783		
Db	513	CAAGTGGAAAG---ACGAAAGCAGGAGAGCGGGGCAAGCTCAACCGCGACGACAGCGACGA 569		
QY	784	CGATCCCGAAGACATGACGCGGACGTGGCACTGAAACGACGAGCGAGGACCGCTGTC 843		
Db	570	CGACGACGACAGAACGACGACGAGTACATCTCTCTCTCGGAGCGAGGACGCGCTGTG 629		
QY	844	GAGGAAGGTGTGATTCGCTCGACGAAAGGTGAACCCGCTACCGGATGGTGTGATCGTGGTGC 903		
Db	630	GAGGAAGGTGCCGATCCGTCGAGCAAGATCAACCCGCTACCGGATCGTGATCGTGCTCCG 689		
QY	904	TCTCGTGTGCTCGCCTTCTTCTCCGCTACCGTATCTCTGACACCCGCTCCGAGACGCAT 963		
Db	690	GCTGGTGGTCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 749		
QY	964	CGGCTGTGGTCTGCTCTCATCTATCTCGAGATCTGTTCCGCTTCCGCTATCTCTGGATCTTCGA 1023		

Db 2459 CTGGTACC-----CTATGGCGCCGCTCAAGTGGCTCGAGCGCTTCGCCCTACCAACAC 2514
Qy 2572 CACCATCTACCCCTTCACTCGCTCCCGCTGCTCGCTACTGCAACCTCTCCCGCGCTG 2631
Db 2515 CATGCTACTACCCCTTCACTCCCTCTCCCTGCTGCTACTGCAACATCCCGCGCTG 2574
Qy 2632 CTCCTCACCGGCAAGTTCATCATGCGTGCATGATGACAGCTTCGCCAGCTCTTCTTCAT 2691
Db 2575 CTTCTCTCACCGGCAAGTTCATCATCCCAAGCTTAACTTTGGCGAGCATATGTTTCAT 2634
Qy 2692 CGCCTCTTCTATGTCATCTTCGGAGCGGATCCTCGAGATCGGTTGGAGCGGGTGAG 2751
Db 2635 AGCGCTTTTCTGTCGATCATCGCGAGCGGGGTGCTGAGCTCGGTTGGAGCGGGTGAG 2694
Qy 2752 CATCGAGAGTGTGGAGGACGAGCATGTTCTGGGTTCATCGCGCGCGTGTCCGCGCATCT 2811
Db 2695 CATCGAGAGTGTGGAGGACGAGCATGTTCTGGGTTCATCGCGCGCGTGTCCGCGCATCT 2754
Qy 2812 CTTGCGCGTGTGAGCGGCTGCTCAAGGCTCTGCGCGGATCGACACCACTTCAAGCT 2871
Db 2755 GTTGGCGGTGTTCAGAGCTCTCTCAAGGCTCTGCGCGGCTGAGACCACTTCAAGCT 2814
Qy 2872 CACTCCAAAGCCACCGGAGGAGGACGAGTTCGAGGTTTCGCGAGCTCTACGCTTCAAGT 2930
Db 2815 GAGTCCAAAGCCCGCGGAGGAGGACGAGTTCGAGGTTTCGCGAGCTCTAACTGTTCAAGT 2874
Qy 2931 GGACACGCTCTCATCCCGCCACCAAGCTGCTCATTAAGTTCATCGCGGTCTGAG 2990
Db 2875 GGACACGCTGCTGCTCGCGGAGGAGGCTGATCATCATCAATGTTGGGAGTCTGCG 2934
Qy 2991 CCGGATCTCGAGCGCATCAACAAAGGTTACAGTCTCGGGGCGCCCTCTTCGGCAAGC 3050
Db 2935 CCGGCGTGTGCGAGCGCGTCAACAAAGGTTACAGGTTCTGGGGCGCCCTGTTGGGAAGC 2994
Qy 3051 TCTTCTCGCTCTGAGTATGTCACCTCTACCGTTTCCCTCAAGGGGCTCATGGGC 3110
Db 2995 TCCTTCTCTCTCTGAGTATGTCACCTCTACCGTTTCCCTCAAGGGGCTCATGGGA 3054
Qy 3111 GCCAGAACAGGACGCCACCGTGTGTTGTCATCTGCTCATCTGCTGCTGCTGCTCATCTTCT 3170
Db 3055 GGCAGAACGAGCGGCCACCAATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCT 3114
Qy 3171 CCTGCTCTGAGTATGAGTACCGTTCATGTCAGGACCAAGGGCGCGAGCTGAGC 3230
Db 3115 CCTGCTCTGAGTATGAGTACCGTTCATGTCAGGACCAAGGGCGCGAGCTGCTCAAGC 3174
Qy 3231 AGTGTGATCATGTTGAGTGTGTT 3256
Db 3175 CATGCGGGGTCTGCTGCTGAGCTGCT 3200

RESULT 13
AF200528
LOCUS AF200528 3745 bp mRNA linear PLN 31-AUG-2000
DEFINITION Zea mays cellulose synthase-4 (Cesa-4) mRNA, complete cds.
ACCESSION AF200528
VERSION AF200528.1 GI:9622879
KEYWORDS .
SOURCE Zea mays
ORGANISM Zea mays
Sukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3745)
Holland,N., Holland,D., Helentjaris,T., Dhugga,K.S.,
Xoconostle-Cazares,B. and Delmer,D.P.
A comparative analysis of the plant cellulose synthase (Cesa) gene family
Plant Physiol. 123 (4), 1313-1324 (2000)
JOURNAL MEDLINE 20398328
PUBMED 10938350
REFERENCE 2 (bases 1 to 3745)

AUTHORS Dhugga,K.S. and Helentjaris,T.G.
TITLE Direct Submission
JOURNAL Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA
FEATURES
source
1. .3745
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
/note="identified and isolated from the maize genome database CDNA collection of Pioneer Hi-Bred International, Inc." .3745
gene
1. .3745
/genes="Cesa-4"
320. .3553
/genes="Cesa-4"
/note="Cesa-4"
/codon_start=1
/product="cellulose synthase-4"
/protein_id="AAF89964.1"
/db_xref="GI:9622880"
/translation="MEGDADGVKSRGGGQVQICGDGVGTTAEGDVFACDVGCP
VCRPCVEYERKDGTOACPOCKTKYKRKGSPIARGEGDDTDADSPNYLASNEDOK
QKADMRWRMNVGSGVGRPKYDSGEIGLTKYDSGEIPRGYIIPSVTNSQISGEIP
GASPDHMMSPPTGNIGRAPFPYVHSPSPREFSGISGNVAKWVRDGMKQDKGT
IPMNTGTSIAPSEGRGVGDIDASTDYNMEDALLDETQPLSRKPLPSPRPINRYRV
IVLRILVLSIFLHYRTNPNRAYPLMLLSVICIWFALSWILDQFPKPPINRETVL
DRLARVDRGEPQSLAAVDIFVSTVDPMKPEPLTANTVLSILAVDYPVKVSCVYS
DGLAAMLTDALEATSEFARKVFPVKYNIERAPRPEWFSOKIDYLDKVKHPSFKVD
RRAMKREYSEPKVRVNLVAKAQKVEEGIMQDGTTPWGNTRDRPHGMLOVFLHSG
KDLTEGNEPLRVLYVSRKRPQFHKKAGAMNALVRSAVLTNRTQTMLENDCHYIN
NSKLEAEMCFLEMDPNLRSVCYVQPFQPDGIDRNDRYANRTVFFDINLRGDGIG
GPVYGTGVFNRTALGYEPTIKQKGGFLSSLCGGRKAKSKSKGSKKSKQKHVD
SSVPVFNLDIIEBEGVAGFDDEKSLMSQMSLEKRFQGSAAFAVASTLMEYGVPOSA
TPESLLKEAHVISCYEDKTEWGTETIGVSVTEDIITGPKMHARGWESIYCMKPR
PAPKAPINLDRNLQVLRWALGSVEILFSPCHPLWYGVGRLKELERPAYINTIY
PLTSLPLLYICILPAICLTGKFIIPISNFASIFWISFISIFATGILBMRWSGVGI
DEWNRNEQFWTIGGISAHLFVAVGILLKVLADINTFTVTSKASDSDGDFAEILMFKM
TLLIIPPTILLINLVGVAFYSAINSGVQSWGFLFKLFFAFWIVHLYPLFLKGLM
GRNQRTPTIVVWMAILLASIFSLWVRIDPFTTRVTGPDQTQCGINC"

ORIGIN
Query Match 40.8%; Score 1398.8; DB 8; Length 3745;
Best Local Similarity 67.2%; Pred. No. 1.3e-174;
Matches 2185; Conservative 1; Mismatches 923; Indels 143; Gaps 9;
Qy 133 GAACGAGTGGTCTGATCCGGGGCCACGAGGACCCCAAGCCGCTCGCGGCGCTGAGCGG 192
Db 307 GAGCTCGTTGCCATGAGGGCGACGGGACGCGTGAAGTCGGGAGGCGCGGTGCGG 366
Qy 193 GCAGGTGTGCGAGATATGCGCGACGAGGTTCGGGCTCACGGTGGAGCGGACCTCTTCGT 252
Db 367 ACAGGTGTGCGAGATCTGCGCGACGCGTGGGACCAACGCGGAGGGGACGCTTTCG 426
Qy 253 CGCTGCAACGAGTGGGCTTCCCGTGTGCGCGCCCTGCTACGAGTACAGCGCCGCGGA 312
Db 427 CGCTCGAGCTGTCGGGGTTTCGGTGTGCGCGCCCTGCTACGAGTACAGCGCAAGGA 486
Qy 313 GGGCAGCGAAGCTGCCCCAGTGAAGCGCTCAAGCGCTCAAGGGGAGCCCGAG 372
Db 487 CGGACGAGGCGTGCCTCCAGTGAAGCAAGTCAAGGCAAGGCGCAAGGGGAGCCCGG 546
Qy 373 GGTTCGGGGAGCATGACGAGGAGGACATCGACGACCTGAGAGCAGAGTTCACATCGA 432
Db 547 GATCCGTGGGGAGGAGGACGACACTGATGCGATAGCGACTTCAATTACCTTGCATC 606
Qy 433 CGACGAGATCAGCAGAGGAGCTGGAGGCAACATGCAGAACAGCCAGATCA----- 485
Db 607 TGGCAATGAGGACCAAGAGCAGAAAGATTGCCGACAAATGCCAGCTGGCGCATGAACGT 666
Qy 486 -----CCGAGGGGATGCTGTCAGCGGAGGATGAGTACGGGAGGGGCCCCG----- 530
Db 667 TGGGGGAGCGGGGATGTTGGTCCGCCAAGTATGACAGTGGCGAGATCGGGCTTACCAA 726

QY 531 --ACAGCGGACGGCAACAACACCCCGAGATCCCGCCCATCATCAACCGGCTCCCGTC 588
Db |||||
QY 727 GTATGACAGTGGCGAGATTCTCTCGGGGATACATCCCATCAGTCACTAAACAGCCAGATCTC 786
Db |||||
QY 589 CGTCCCGTGAGCGGTGAGTTCC--GATTACCAACGGGTATGGCCACGGCGAGGTCTCGT 647
Db |||||
QY 787 AGGAGAAATCCGTGTGTCTCCCTGACCATCATATGATGCCCAATCGGAAATGG 846
Db |||||
QY 648 CTTCCCTGACAAAGCGCATCCATCCGTACCTGTCTTGAGCAGGGAGTGCCTAGTGGG 707
Db CAAGCGTCTCCATTTCC--CTATGTGAACCATTCGCCAAATCCGTCAAGGGAGTTCTCTG 905
QY 708 ACGAGAAAGAGTGAAGTGAAGAGAGAGATGGACGACTGGAAGTCCAGACGGGCA 767
Db |||||
QY 906 GTAGCATTTGGGAATGTTCCCTGGAAAGAGAGGGTTGATGGCTGGAAATGAAGCAGGACA 965
QY 768 -----TCC 770
Db AGGGAGCATTTCCCATGACGAATGGCAACAGCATTTGCTCCCTCTGAGGGTGGGGTGTG 1025
QY 771 TCGCGCGCGCGCGGATCCGGAAGACATGGACGCGAGCTGGCACTGAACGACGAGGCGA 830
Db GTGATATTGATGCAATCACTGATTACACATGGAAGATGCCTTATTGAACGACGAACTC 1085
QY 831 GGCAGCCGCTGTGAGGAAAGTGTGATCGCGTGCAGCAAGGTGAACCCGTTACCGGATGG 890
Db |||||
QY 1086 GACAGCCTCTATCTAGGAAAGTTCCACTTCTCTCCAGGATAAATCCATACAGGATGG 1145
QY 891 TGATCGTGGTGGTCTCGTGTGCTCGGCTTCTTCTCCGTGCGTATCTGCACCCG 950
Db TCAITGTGCTGATGATTGTTCTTAAGCATCTCTTGCACTACCGGTATCAAAATCCTG 1205
QY 951 TCCCGGAGCCATCGGGCTGTGGCTGTCTCATCATCTGCAGATCTGGTTTCGCCATCT 1010
Db TGCGCAATGCATACCAATATGGCTTCTATCTGTATATGTGAGATCTGGTTTGTCTCTT 1265
QY 1011 CCTGATCCTCGACAGTTCCTCCAGTGGTTCCTCCATCGACCGGAGAGCTACCTCGACC 1070
Db CGTGGATATTGGATCAGTTCCTCAAGTGGTTCCTCAATCAACCGGAGAGCTACCTTGATA 1325
QY 1071 GCCTCTCCTCAGGTACGAGAGGAGGAGCGCTGCTGCTGTCCGGGTGACCTGT 1130
Db GCGTGGCATTAAGGTATGACCGGAAAGTGAAGCATCTCAGTTGGCTGCTGTGACATTT 1385
QY 1131 TCGTGACGACGGTGCACCGCTCAAGGAGCGCGCTGGTGCACCGCCAAACACCGTCTCT 1190
Db TCGTCAGTACAGTCGACCCCAATGAAGAGCTCTCTTGTCTACCTGCCAATACCGTGTAT 1445
QY 1191 CCATCTCGCGTACACTACCCCGTGGCAAGGTCTCTGCTACGTCTTCGACGACGCG 1250
Db CCAATCTGCTGTGGATTACCTGTGGATAAGGTCTCTTGTCTATGTATCTGATGGAG 1505
QY 1251 CGTGCATGCTCAGCTTCAGTTCGTGTGAGAGCGGCGAGTTCGCGCGCAAGTGGTGC 1310
Db CTGCGATGCTGACATTTGATGCACGTAGCTGAGACTTCAGAGTTTGTAGAAATAGGGTAC 1565
QY 1311 CCTTCTGCAAGAAGTTCCGGCATCGAGCCCGCGCGCGCGAGTTCCTACTTCTCGCTCAAG 1370
Db CATTTGTTAAGAGTACAAATGNAACCTAGAGTCTCTGATGGTACTTCTCCAGAAA 1625
QY 1371 TCGACTACCTCAAGGACAAAGTGCAGCCCACTTGTGTGAGAGCGCGCGCCCATGAAGA 1430
Db TTGATTAATTGAAGACAAAGTGCACCTTCATTTGTTTAAAGACCGCGCGGCCATGAAGA 1685
QY 1431 GAGAGTATGAGAGTTCAAGTTCGGATCAACGGCTGCTGGCCAGGCGCATGAAGTGC 1490
Db GAGAAATATGAAGAAATTCAGAGTTAGGGTAAATGGCTTTGTGCTAAAGGACAGAAAGTTC 1745
QY 1491 CGGACAGGGGTGATCATGAAGACGCGACCGCTGCGCGGGAACAACACCGCGACC 1550
Db CTGAGGAAGGATGATCATGCAAGATGGCACACCATGCGCCAGGAAACATACCGGGACC 1805

QY 1551 ACCCGCGCATGATCAAGGTGTTCTTGGGCACACAGCGCGGCCACGACACCCAGGGAACACG 1610
Db ATCTGTGAATGATTGAGTTTTCAGGTTTCTTGTGTGTCACAGTGTGGCTTGATACAGGGAATG 1865
QY 1611 AGCTGCCCGCTCGTGTGATGCTCTCCGTGAGAGCGCCCGGATTCAGACACACACAGA 1670
Db AGCTACCCCGTTTGGTCTATGTTCTTGTGAAAAGCGCTCTTGGATTCAGCATCAAGA 1925
QY 1671 AGGCGGCGCCATGAACGCTCTGATTCGCGTCTCCGCGTGTGACCAACCGCGCATTC 1730
Db AAGCTGGTGCCATGAATGCTCTTCTGTCTCAGCTGTCTTACCAATGACAATACA 1985
QY 1731 TGCTCAACTTGGACTGTGATCACTACATCAACAACAGCAAGGCCATCCGGGAGGCGATGT 1790
Db TGTGAAATCTTGAATGTGATCACTACATTAACAACAGTAAGGCTCTCAGGGAAGCTATGT 2045
QY 1791 GTTTCCTCATGAGACCTCTCAGTCCGCGGAGAGTCTGCTAGCTCAGTTCAGTTCGCGAGAGT 1850
Db GTTTCCTTATGACCCCTAACTAGAAAGAGTGTCTGCTAGCTCAGTTCCTCCAGAT 2105
QY 1851 TCGACGCGATCGAGCTGCACGACCGATACGCTAAACAGGAAACACCGTCTTCTTCGACATCA 1910
Db TCGATGCAATTCGACAGGAATGATCGATATGCCAACAGGAAACACCGTGTTCGATATTA 2165
QY 1911 ACATGAAGGGCTCGACGGCATCCAGGCCCGGTGTAGTCTGGACAGGCTGCTGTTC 1970
Db ACTTGAGAGGTCTTGTATGCGCATCCAGGACCAAGTTTATGTGCGAACTGGCTGTGTTTCA 2225
QY 1971 GGGCCAGGCGCTCTAGGCTACAACTCCCAAGGACCAAGA----- 2015
Db ACCGAAACAGCTCTATATGTTTATGAGCCCAATTAAGCAAGAGAGGTGGTTCTTGT 2285
QY 2016 -----GGCCCAAGATGTGACCTGCGACTGTGCTGCCGTGCTTCGCGCGCAAGA 2063
Db CATCACTATGTGGCGTGAAGAGCAAGCAATCAAGAGAGGCTCGACACAAGA 2345
QY 2064 AGCGGAAACAGCCCAAGAC-----GGGCTGCCGAGG 2096
Db AGTGCAGAAAGCATGTGACAGATTTCTGTGCCAGTATTCAACCTTTGAAGATATAGAGGAG 2405
QY 2097 GCACCGCTGATATGGGAGTAGATAGCACAGGAGATGCTCATGTCCACATGAATCTCG 2156
Db GAGTTGAAGGGCTGGATTTGAACAGGAAATCACTTCTTATGTCTCAATAGACCTGG 2465
QY 2157 AGAAGCGGTTCGGGACGTCCGCGCTTCGTGCTGACGCTGATGAGGAGGAGCGCG 2216
Db AGAAGATTTGGCCAGTCCGCGAGTTTGTGCTCCACTCTGATGGAGTATGGTGTG 2525
QY 2217 TCCCTCCTGTCGAGCCCGCGCTCCTCAAGGAGGCGCATCCATGTTCATGCTAGCTGCG 2276
Db TTCTCAGTCCGCAACTCCGAGTCTCTTCTGAAAGAGCTATCCATGTTATAAGCTGTG 2585
QY 2277 GCTACGAGGACAGACCGACTGGGGCTGGAGCTGGGGTGGATCTACGGTTCGATCAAG 2336
Db GCTATGAGGACAGACTGGAATGGGAACTGAGATCGGGTGGATCTACGGTTCGTGACAG 2645
QY 2337 AGGACATCTGACCGGGTTCAAGATGCACTGCGCGGGTGGCGTCTCGTGTACTGCTATGC 2396
Db AAGACATCTCACCGGATTCAGATGCAACGCGCGAGGCTGGCGGTGATCTACTGCTATGC 2705
QY 2397 CGAAGCGGGCGGCTTCAAGGGTTCGGCGCGGATCAATCTATCGGACCGCTTCACACGAG 2456
Db CCAAGCGCCAGCTTCAAGGGTCTGCGCCCATCAATCTTTCGAGCCGCTCGTGAACCAAG 2765
QY 2457 TGCTCCGCTGGCGCTGGGCTCGCTGAGATCTTCTCAGCGGACAGCCCGCTGCTGT 2516
Db TGCTCCGCTGGGCTCTTGGGTCCGTGAGATCTCTTTCAGCCGCGACTGCCCGCTGTGTGT 2825
QY 2517 ACGGCTTACAAGAACCGCAACTCAAGTGGGTGAGCGCTTCGCTTACATCAACCAACCA 2576
Db ACGGCTTAC---GGAGGGCGGCTCAAGTTCCTGGAGAGATTCGCGTATCAACCAACCA 2882
QY 2577 TCTACCCCTTCAACCTCGCTCCCGTGTCTGCTACTGCAACCTCCCGCGCGTCTGCTCTCC 2636

```

2883 TCTACCGCTCAGTTCATCCGCTTCTCTACTGCTCATCTGCGCGCCATCTGTCTGC 2942
Qy
2637 TCACCGGCAAGTTCATCATCGCGTCGATTAGCAGTTTCGCGAGCTCTTCTTCATCGCCC 2696
Db
2943 TCACCGGAAAGTTCATCTCCAGAGATCAGCACTTCGCGAGCATCTGTTTATCTCCC 3002
Qy
2697 TCTTATGTCATCTTCGCGACGGGCATCTCGGAGATCGCGTGGAGCGGGTGAGCATCG 2756
Db
3003 TCTTATGTCATCTTCGCGACGGGCATCTCGGAGATGAGTGGAGCGGGTGCGCATCG 3062
Qy
2757 AGGAGTGTGGAGGAACGAGAGTTCTGGGTTCATCGCGCGGTGTCGCGCATCTCTTCG 2816
Db
3063 ACAGAGTGTGGAGGAACGAGAGTTCTGGGTTCATCGCGCGGTGTCGCGCATCTCTTCG 3122
Qy
2817 CCGTGTGTGAGGGCTCTCTCAAGTCTCTCGCGGATCGACCACTTCACCGTCACTT 2876
Db
3123 CCGTGTGTGAGGGCTCTCTCAAGTCTCTCGCGGATCGACCACTTCACCGTCACTT 3182
Qy
2877 CCAAGGCCACCGGCGACGAGGACGACGAGTTTCGCGGAGCTCTAGCGCTTCAAGTGGACCA 2936
Db
3183 CCAAGGCCCTCG--GACGAGGACGGGACTTTCGCGGAGCTGTACATGTTCAAGTGGACGA 3239
Qy
2937 CGTCTCTCATCCCGCCACACAGCTGTCTCATCTTAACGTCATCGCGGTCTGTCGCGGCA 2996
Db
3240 CGTCTCTCATCCCGCCACACAGCTGTCTCATCTTAACGTCATCGCGGTCTGTCGCGGCA 3299
Qy
2997 TCTCCGACGGCATCAACAAACGGGTACAGTCTCTGGGGGCGCTCTTCGCGCAAGCTCTTCT 3056
Db
3300 TCTCTACGCGATCAACAGCGGATACAGTCTGTCGCGGCGCTCTTCGCGCAAGCTCTTCT 3359
Qy
3057 TCGCTCTTCTGGTTCATCGTCACTCTACCGCTTCTCAAGGGGCTCATGGGGCGCCAGA 3116
Db
3360 TCGCTCTTCTGGTTCATCGTCACTCTACCGCTTCTCAAGGGGCTCATGGGGCGCCAGA 3419
Qy
3117 ACAGGACGCCACCGTTGTTGTCATCTGTTGTCATCTGTCGCTCATCTTCTTCTTCTTCT 3176
Db
3420 ACCGACCCCGACCATCGTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 3479
Qy
3177 TCTGGGTTCAGATCGACCTTCTCATCTCAGGACCAAGGGCGCGGAGCTCAGGAGTGTG 3236
Db
3480 TGTGGGTTCGATCGACCTTCTACACCGCGCTCACTGCGCGGATACCAACAGTGTG 3339
Qy
3237 GCATCAATGCT 3248
Db
3540 GCATCAATGCT 3551
```

```

RESULT 14
AY639654
LOCUS
DEFINITION
Pinus radiata cellulose synthase catalytic subunit (CesA1) mRNA,
complete cds.
ACCESSION
AY639654
VERSION
AY639654.1 GI:49182339
KEYWORDS
Pinus radiata (Monterey pine)
ORGANISM
Pinus radiata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
1. (bases 1 to 3911)
Klein,A.S., Tibbitts,J., Steven,R. and Anthony,B.
Cellulose synthase genes in Conifers: what we know and what we need
to learn
(in) Havashi, T. (Ed.);
PLANT CELL WALLS;
(2005) In press
2. (bases 1 to 3911)
Klein,A.S., Tibbitts,J., Steven,R. and Anthony,B.
Direct Submission
Submitted (27-MAY-2004) Biochemistry and Molecular Biology,
University of New Hampshire, 46 College Road, Durham, NH 03824, USA
Location/Qualifiers
```

```

source
1..3911
/organism="Pinus radiata"
/mol_type="mRNA"
/db_xref="taxon:3347"
gene
1..3911
/genes="CesA1"
/notes="PrCesA1"
5'UTR
1..168
/genes="CesA1"
169..3423
CDS
/genes="CesA1"
/EC_number="2.4.1.12"
/notes="glycosyltransferase"
/codon_start=1
/product="cellulose synthase catalytic subunit"
/protein_id="AAT57672.1"
/db_xref="GI:49182340"
/translacion="MEASGLVAGSHNRNEFVVIHGHEGPKPLNTLSGHVCQICGEDV
GLNTDGEFLVACNECFVPCRYEYEREGNQSCPCNTRYKRGSRPVEGDDDEE
DVIDIEHFNVAQRNQIITEAMLRMSYGRGPDNDENSIQAHNPPELPPIQIPLVAN
GHSVSGEIPTSYVADNQLAMPALKRHPSEPGSGRIIMDPNMDIAGSFGNVSW
KERGDKSKENKSGOLDMTEGRYOYNGGFAPNEPEDYIDPMPMTDEARQPLSRKVP
IPSSKINPYRMVIVIRLIVLIGIFLRYLLNPKNAYGLMATSIVICWPAISWILQF
PKWLPISRTYLDRLSYREREGEPMAPVDLFSVTDPLKEPPLVANTANTLISLV
DYPVDNVSYSDSGASMLTFSELSSETSFARKWPFCKFDIEPRAPETIYFSQIDY
LKDQFQTTVKERRAMKREYEBFKVIRNLVAKASKVPKQEGMTQDGTWPQNNTRDH
PGMIOVFLHSGGLDTEGNELPRLVYVREKRPFOHKKAGAMNALVRSAVLTNAP
FMNLDCDHVINNSKAIREGMCFMDPOVGRKVCYVQFQRPEDGIDRNDRYANRNTVF
FDINMGLDGIQSPVYVGTGCMFRQALYGYGPKPKPKMVTCDCLPCCGPKRKP
KNSSKSGISAPAPAYNDGIEBEGVEDDRLALMSQDLEKKFGQSSAFVOSTLME
NGVQPTANPAELLKEAHIHISGVEDKTEWGLGMIYGSVTEIDLTGPKMTGRMR
SYICMPKRAAFKSGAPINLSRLNQLRWALASVEIFMSRHCFPIATFISIFATILE
AWINTVYPTSLPLATCTLPASVLLTKGFVLPQISTFASLFIATFISIFATILE
MWSGVSIIEWMNEQFVIGVSAHFRAVIOGLLKLVLADIDNTFTVTAKASDGBRG
ELIYAFKWTLLIPPTLLINLVGVVGVADAINNGFQSNGLPLGLKLPFAPFWIVHLY
PELKGUMGQNRPTTIVIVISILLASVFSFLVWRIDIFLSKVAGPDTKQCGINC"
3'UTR
3424..3879
/genes="CesA1"
ORIGIN
Query Match 39.8%; Score 1371.8; DB 8; Length 3911;
Best Local Similarity 66.8%; Pred. No. 4.5e-171; Indels 166; Gaps 10;
Matches 2200; Conservative 0; Mismatches 937;
Qy 90 CGATGAGGCCAGCGCGGCTGCTGCGCGGCTCGCACAAACCGGAAACGAGCTGCTGCTGA 149
Db 167 CAATGGAAGCCAGCGCGGCTGTTGCCGTTCTCTAACAACAGAGTCTGCTGCTCA 226
Qy 150 TCCGGGGCCACAGGACCCCAAGCCGCTGCGGGCGCTGAGCGGGCGAGGTGTGCGAGATAT 209
Db 227 TCCATGGACATGAGGGCCGAGCCTTTGAACACGTTGAGTGGCCACGCTCTGCCAGATTT 286
Qy 210 GCGGCGACGAGTTCGGGCTCAAGTTCGAGCGGCGACCTCTTCGTCGCTGCAACGAGTGGC 269
Db 287 GTGGCGAGGACGTCGGGCTTAACACAGAGCGCGAGCTGTTCTGTTGCTGTAATGAGTGGC 346
Qy 270 GCTTCCCGTGTGTCGGGCTCTGTACAGTACAGCGCCGGGAGGCGCACGCAACTGCC 329
Db 347 GGTTCCTCTGTCGCGCGCTGTATGAGTACAGAGAGAGAGAGAAATCATGCTGCTGCC 406
Qy 330 CCCAGTGCAAGACGCGCTACAAGCGCTCAAGCGGAGCCGAGGGTTCGCGGGAGCAATG 389
Db 407 CGCAGTGCAATCTCGTTACAAGCGTCAAAAAGGGAGTCCACGGGTGGAAGGTGACCATG 466
Qy 390 ACGAGGAGGACATCGACGACCTGGAGCAGAGTTCAACATCGACGAGGAATCAGCAGA 449
Db 467 ATGAAGAAGACGTTGATGATAGATAGAAATTAATTTGAGAGGCTCAGCAAGAAACA 526
Qy 450 GGCAGCTGGAGGCAACATGACAGAACAGCGAGATCACCGAGGCGATGCTCCACGGGAGGA 509
Db 527 GGCAG-----CAGATCACCGAGGCGATGCTCCACGGACGCA 562
Qy 510 TGAGCTACGGGAGGGGCCCGGACGAGCGGCGACGGGCAACACACCCCGCAGATCCCGCCCA 569
```


QY	2599	GCTGCTCGCCTACTGCACCCCTCCCGCGCTGCTGCTCCTCACCGGCAAGTTCATCATGCC	2658
Db	2778	ACTCATTTGCTATTGCACACTTCACGCGCTCAGTTTGTCTCAGTGGCAAAATTTGTGATCCC	2837
QY	2659	GTGCTATTAGACGCTTCGCGAGCTCTCTTCATCGCCCTTCTCATCTCCATCTCCGAC	2718
Db	2838	TCAGATCAGTACTTTTGCAGTCTATTTTATAGCTCTTTTCTCATCTCAATTTTGGCCAC	2897
QY	2719	GGGCATCTCGAGATCGGTGGAGCGGGTGAGCATCGAGAGTGGTGGAGGAACGAGCA	2778
Db	2898	TGGTATCTTGGAAATGAGGTGGAGTGAGTGAACATTTAGATTTGGCGAAATGAACA	2957
QY	2779	GTTCTGGGTCAATCGCGCGGTGTCCGCGCATCTCTTGGCGCTCGTCAGGCGCTGCTCAA	2838
Db	2958	GTTCTGGGTATTGGAGGGGTTTCTGCACATTTTGTGCAATTTTTCAGGTCTGCTCAA	3017
QY	2839	GGTCTCGCGGATCGACACCACTTCAACCTCAGCTCCAGGCAACCGGCGACGAGA	2898
Db	3018	GGTACTGGCAGGCAITGATACAAATTTTCAGTCACTGCGCAAGGCATC-----AGATGA	3071
QY	2899	CGACGAGTTTCGCGAGCTCTACGCTTCAAGTGGACACAGCTCCTCATCCCGCCACCAC	2958
Db	3072	CGGTGAGTTTGGGAACTGTATGCAATTCAAATGGACCACTCTCTATTCTCTCTACAC	3131
QY	2959	GCTGCTCATATTAACTCATCGGCGTGTGCGCGCATCTCCGACGCCATCAACAACGG	3018
Db	3132	CCTGCTTGTTCATCAACTTGTGGGGTGGTGTGGCGTAGCAGATGCAATCAACAATGG	3191
QY	3019	GTACCACTCTGGGGGCCCTCTTCGGCAGCTCTTTCGCCCTTCTGGGTCTATCGTCCA	3078
Db	3192	ATTTAGTCAATGGGTCCTCTCTTGGGTAAAGCTTTTCTTGGTCAATTCGGGTCAATTTGCA	3251
QY	3079	CCTCTACCGGTTCTCAAGGGGTCATGGGGCGCCAGAACAGGACGCCACCGTGTGTGT	3138
Db	3252	CCTGTATCTTCTCTCAAGGTTCTCATGGGCGAGAACCAACACCACTCTGTGT	3311
QY	3139	CATCTGGTCAATTCGTGCGCTTCCATCTTCTCCCTGCTCTGGGTGAGATGACCCCTTT	3198
Db	3312	TATTTGGTCAATTCGTGCACTCTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3371
QY	3199	CATCTGTCAGAACCAAGGGCCGACGTCAGGCTGTCGATGTCGATCAATTCGTGCTTTTA	3258
Db	3372	CTTGAGTAAGTTTAAAGGCCAGATATAACAATGTGGCATCACTGCTGATTTCTTTT	3431
QY	3259	TTA 3261	
Db	3432	ATA 3434	
RESULT 15			
BD236020			
LOCUS			
DEFINITION			
Materials and method for modification of plant cell wall			
polyaccharides.			
ACCESSION			
BD236020			
VERSION			
BD236020.1 GI:33045790			
KEYWORDS			
JP 2002527056-A/42.			
SOURCE			
Pinus radiata (Monterey pine)			
ORGANISM			
Pinus radiata			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.			
REFERENCE			
1 (bases 1 to 3851)			
Blobeberg, L.N.			
Materials and method for modification of plant cell wall			
Patent: JP 2002527056-A 42 27-AUG-2002;			
JOURNAL			
GENESIS RESEARCH AND DEVELOPMENT CORP LTD, FLETCHER CHALLENGE			
FORESTS LTD			
COMMENT			
OS Pinus radiata (radiata pine)			
PN JP 2002527056-A/42			
PD 27-AUG-2002			
PP 08-OCT-1999 JP 2000575985			
PR 13-OCT-1998 US 09/170862, 11-AUG-1999 US 60/148426 PI			

LEONARD NATHAN BLOKSBERG			
PC C12N15/09, A01H5/00, C12N5/10, C12N9/00, C12N9/12, C12N15/00, C12N5/			
PC 00			
CC Materials and method for modification of plant cell wall CC			
polyaccharides			
PH	Key	Location/Qualifiers	
FT	source	1. .3851	
/organism='Pinus radiata (radiata pine)'			
/mol_type='genomic DNA'			
/db_xref='taxon:3347'			
FEATURES			
source			
1. .3851			
/organism='Pinus radiata'			
/mol_type='genomic DNA'			
/db_xref='taxon:3347'			
ORIGIN			
Query Match 39.8%; Score 1368.6; DB 6; Length 3851;			
Best Local Similarity 66.5%; Pred. No. 1.2e-170;			
Matches 2198; Conservative 0; Mismatches 939; Indels 166; Gaps 10;			
QY	90	CGATGGAGGCCAGCGCCGGGCTGGTGGCCGGCTCGCACAAACCGGAACGAGCTGGTGTGA	149
Db	151	CAATGGAAGCCAGCGCGGCTTGGTTGCCGTTCTCATACAGAAACGAGTTCGTGTCA	210
QY	150	TCCGGGGCCACAGGAGACCCCAAGCCCTGCGGGCGCTGAGCGGCGAGGTGTGCGAGTAT	209
Db	211	TCCATGGACATGAGGAGCCGAAGCCCTTGAACACGTTGAGTGGCCACGCTCTGCCAGATT	270
QY	210	GCGGGCACGAGGTTCGGGCTCACGGTGGACGGCGACCTCTTCGTGCGCTGCAACGAGTGGC	269
Db	271	GTGGCGAGGACGTCGGGCTTAAACAGACGGCGAGCTGTTGTTGCTGTATAGTGTGCG	330
QY	270	GCTTCCCTGTGTGCCGCGCTCTACGAGTACGAGCGCCGGAGGGACACGCAAACTGCC	329
Db	331	GGTTTCCTGTCTGTGCGCGTGTATAGTACGAGAGCGAAGAAATCAGTCGTGTC	390
QY	330	CCCAGTGCAGACGCGCTACAGCGCTCAAGGGAGCCGAGGCTTCCCGGGACGATG	389
Db	391	CGCAGTGCATTAATCTCGTTTCAAGCGTCAAAAAGGGAGTCCACGGGTGGAAGGTGACGATG	450
QY	390	ACGAGGAGGACATCGACGACCTGGAGCACGAGTTCAACATCGACGACGAAATCAGCAGA	449
Db	451	ATGAGAGAGCGTTGATGACATAGAACATGAATTTAATGTGGAGATCAGCAAGAACA	510
QY	450	GGCAGCTGGAGGGCAACATCGAGAACGCGCAGATCACCGAGGCGATGCTGCACGGCAGGA	509
Db	511	GGCAG-----CAGATCACCGAGGCGATGCTCCACGGACGA	546
QY	510	TGAGCTACGGAGGGGCGCCGACGAGCGGCGAAGGCAACACCCCGCAGATCCCGCCCA	569
Db	547	TGAGCTATGGCCGAGGTCCCGACGACGAGAAATTCGCAGATTGCTC--ATAATCCAGAGCT	604
QY	570	TCATCACCGGCTCCCGCTCCGTGC-----CGGTGAGCGGTGAGTTTCCGAT	615
Db	605	TCTCCGAGATTCCTGTACTTGCACACCGCCACTCGGTTGTGAGTGGGAGATTCCAAC	664
QY	616	TAC---CAACGGGTATGGCCACGGCGAGGTCTCGTCTTCCCTGCACAAAGCGCATCCATCC	672
Db	665	GTCACTACTACGACAGCAACCAATTTGTTGCAACCTGCAATGCTGAAGCGTGTGATCC	724
QY	673	GTACCTCTGTCTGAGCCAGGGAGTCCCAAGTGGGACGAGAGAAAG-----719	
Db	725	AAGTCCGAGCCGGGAGTGGAGGATCATCATGATTCACAAACAGGATATTGGTTCTTA	784
QY	720	-----AAGTGAAGTGGAGGAGGATGGACGACTGGAAGTCCAAAGCAGAGGCGAT	768
Db	785	TGGCTTTGGGAAGTGTCTTGGAGAGGCGAGCGGATGTTTATAATCGAAGGAAACAA	844
QY	769	CCTCGCGCGCGCGCGATCCCGAAG-----794	
Db	845	ATCAGGCCAGTTGGATATGACGGAAGGGAGATATCAATATAATGGGGGTTTGCACCAA	904
QY	795	-----ACATGACGCCCGCGCGATCCCGAAG-----840	

Db 905 TGAGCCTGAAGATTATATTGATCCGATATGCCAATGACCGATGAACGAGGACGCCACT 964
Qy 841 GTCAGGAAGGTGTCGATCGCGTCGAGCAAGGTGAACCCGTAACCGGATGATGTCGTGGT 900
Db 965 GTCCGAAAAGTGCCAATTCCTTCAAGCAAAATAAATCAATACCGAATGATGTCATGTAAT 1024
Qy 901 GCGTCTGTTGTCGTGCTCTTCCCTCCGATACCGTATCTGTCACACCCGTCGCGGACGC 960
Db 1025 TCGACTGATAGTGTGGGTATTTTCTCCGCTATCGTCTCTGAAATCCAGTGAAGAATGC 1084
Qy 961 CATCGGCTGTGGCTGCTCTCCATCATCTGCGAGATCTGGTTTGCCCATCTCTCGGATCCT 1020
Db 1085 ATATGGGCTCTGGGCCACTTCTATCGTTTGTGAAATCTGGTTTGCTTGTCAATGATCT 1144
Qy 1021 CGACAGTCTCCCAAGTGGTTCCCATCGACCGCGAGACGTACTCGACCGCTCTCCCT 1080
Db 1145 TGATCAGTCTCCCAAGTGGTTGCTTATCAGTCGTGTAACGCTATCTTGATCGACTGTCAAT 1204
Qy 1081 CAGGTACGAGGGAAGGGAGCGCTCGCTGCTGTCGCGCGTGACCTGTTGCTGAGCAC 1140
Db 1205 AAGGTACGAGGAGGCGCAACCATCAATGCTTGTGCACCTGTTGACCTCTTGTGAGTAC 1264
Qy 1141 GGTGGACCCGCTCAAGGAGCGCGCTGGTGACCGCCAAACACCGTGTCTCCATCTCGC 1200
Db 1265 TGTAGATCCACTGAAGGAGCTCTTGTGTTACTTGCCCAATACAGTATTAATCAATCTTTC 1324
Qy 1201 CGTAGACTACCCGTGACAAAGTCTCTCTGTACTGCTTCGACGACCGCGCTCGATGCT 1260
Db 1325 AGTAGACTACCTGTAGACAATGTCTCTGTATGTCTCTGATGACGAGCGCTCGATGCT 1384
Qy 1261 GAGCTTCGAGTCTGTGCGAGCGCGGAGTTCGCGCGCAAGTGGGTGCGCTTCTGCAA 1320
Db 1385 TACTTTTGAATCTCTCTGAGACTCAGAAATTTGCCAGAAATGGGTGACATTTCTGCAA 1444
Qy 1321 GAAGTTCGGCATGAGCCCGCGCGCGAGTCTACTTCTCGCTCAAGGTGCGACTACCT 1380
Db 1445 GAAATTCGACATGAGCCTCGCGCTCCGGAATCTATTCTCTCAGAAATTTGACTATCT 1504
Qy 1381 CAAGGACAAGGTGACGCCACTTTGTCGAGAGCGCGCGCCATGAAGAGAGATGA 1440
Db 1505 GAAGGACAAATTTCAACCCACTTTGTCAAGAGCGCGTGCCATGAAGAGAGAAATGA 1564
Qy 1441 GGAGTTCAGGTCCGGATCAACGCGCTGGTGGCCAAAGCCCATGAGGTGCGGCGAGGG 1500
Db 1565 AGAATTCAGGTGCGCATCAATCGTTGGTTGGTGAAGAGCGCTCTAAAGTGCCCAAGGAAG 1624
Qy 1501 GTGATCATGAAGACGCGCGCTGGCCCGGAAACAAACCCGCGACCAACCCCGCAT 1560
Db 1625 ATGGACAAATGAACAGCGGTACGCTTGGCTGGTAAATATACCCGTGACCAATCTCGTAT 1684
Qy 1561 GATCCAGTGTCTCTGGGCCACAGCGCGGCCACGACACCGAGGGCAACGAGCTGCCCG 1620
Db 1685 GATCCAAGTGTCTTGGGTACAGTGGCGCTCGATACAGAGGCAATGAGCTTCTCTCG 1744
Qy 1621 CCTCGTGTAGTCTCCCGTGAAGCGCCCGGATTTCCAGACACAAAGAGGCGCGCGC 1680
Db 1745 GCTAGTATATGTTTCTCGTGAAGAGACCTGGTTTCCAGCATCACAAAGAGCGCGTGC 1804
Qy 1681 CATGAACGCTCTGATTCGCTCTCCGCTGTGACCAACGCGCATTCATGCTCAACTT 1740
Db 1805 CATGAATGCTTGGTTTGGGTTTCTGCTGTGCTACCAATGCTCCATTTATGCTGAATCT 1864
Qy 1741 GGACTGTGATCACTACATCAACAAACAGCAAGGCCATCCGGAGGCCATGTGCTTCTCAT 1800
Db 1865 GGATTTGTGATCACTACATTAACATAGCAAGGCAATCAGGGAAGGCGATGCTTTTATGAT 1924
Qy 1801 GGACCTCAGTTCGGCCGGAAGTCTGCTAGTTCAGTTCGCGAGAGGTTTCAGCGCAT 1860
Db 1925 GGATCTCCTCAGGTTGGGGAAGGTTCTGTTATGTCCTCAATTTCCCTCAGAGATTCGATGAT 1984
Qy 1861 CGAGCTGCACGACCGATACGTAACAGGAACACCGTCTCTTCGACATCAACATGAAGGG 1920
Db 1985 TGATCGCAATGACCGTTACGCCAATCGAAACACCGTATTTCTTTGATATCAACATGAAGG 2044

Qy 1921 GCTGGACGGCATTCAGGCCCGGTGTACGTCCGGACAGGGTGCCTGTTTCGGCGCCGACGC 1980
Db 2045 TCTGGATGGAATTCAGAGGCGCTGTATATGTGGGAATCGATGCTGTTCAGAGACAAGC 2104
Qy 1981 GCTCTACGGCTACAAACCTCCCAAGGACCCCAAGAGGCCCAAGATGTTGACCTCGGACTG 2040
Db 2105 TCTATATGGGTATGGGCTCCCAAGGCCCAAGCTCCCAAGATGGTGACCTGTGATTG 2164
Qy 2041 CTCGCCGTGCTTCGGCCGCAAGAG----- 2065
Db 2165 TCTCCCTTGTTCGGTCTCTGTAAGAGTCTCCGAAGAAAATAGTAGCAAGAAAAGTGC 2224
Qy 2066 -CGAAACACCGCAAGAGCGGGTCCCGAGGAGCACCGCTGATATGGAGTA-----GA 2118
Db 2225 AGGAATCCAGCTCCCGCTTACAATCTGAGCGGGATCGAGGAAGAGTAGAAGGTTATGA 2284
Qy 2119 TAGCGAAGAGATGCTCATGTCCACATGAATCTCGAAGCGGTTTCGSGCAGTCCGC 2178
Db 2285 TGACGAAGAGCATTTGTTGATGAGCCAACTAGACTTCGAGAAGAAATTTGGCCAGTCTTC 2344
Qy 2179 GGGTTCGTCAACGTCCGACGTGATGGAAGAGCGGCTCCCTCTTCGTGAGCCCGC 2238
Db 2345 AGCTTTGTTCAATCCACTCTGATGGAATGGTGTGTTCCGCAACAGCAATCCAGC 2404
Qy 2239 CGGCTCCTCAAGAGGCCATTCATGTATCAGCTCGGGCTACGAGGACAAAGCCGACTG 2298
Db 2405 TGAATTTGTAAGAGGCTATTCTATGTCATCAGCTGTGGATATGAAGACAAAACGGAATG 2464
Qy 2299 GGGCTGAGCTGGGTGGATCTACGGTGCATCACGAGGACATCTGACGGGTTCAA 2358
Db 2465 GGGAAAGAGCTGGATGGATCTATGGATCAGTCACAGAGACATCTGACTGATTCAA 2524
Qy 2359 GATGCACTGCGGGGTGGGCTCCGTGTACTGATGCCAAGCGGCGGCTTCAAGGG 2418
Db 2525 GATGCACTCGAGGCTGGCGTCCATTTACTGTATGCCCAACAGCAGCATTCAGAGG 2584
Qy 2419 GTGCGGCCCATCAATCTATCGGACCGTCTCAACAGGCTGCTCGGTGGGCGCTGGGGTC 2478
Db 2585 GTCTGCTCCAATCAATCTATCAGACCGTTTGAACACAGGTGTGCGTTGGGCTTGGGATC 2644
Qy 2479 CGTGGAGATCTTCTCAGCGCGCACACCCCTGCTGTAGCGCTTACAAGAACGCAACCT 2538
Db 2645 AGTAGAAATTTTCATGAGCAGACATTCGCCCAATCTGATGCTATGCG---GGAGGTCT 2701
Qy 2539 CAAGTGGCTGGAGCGCTTCGCTTACATCAACACCAACCATACCCCTTACCTCGCTCCC 2598
Db 2702 GAAATGGCTGAAGATTTGCTTATCAACACCAATGTCTATCCATTCACCTCTCTTCC 2761
Qy 2599 GCTGCTGCTTACTGCAACCTCCCGCGCTGCTCTCACCGGCAAGTTTCAATCATGCCC 2658
Db 2762 ACTCATTTGCTTATGACACTTCAGCGCTCAGTTTGTGCTCACTGGCAAAATTTGTGATCCC 2821
Qy 2659 GTCGATTAGACCTTCGCGAGCTCTTCTTCAATCGCCCTTCTCATGTCTCATCTTCGCGAC 2718
Db 2822 TCAGATCAGTACTTTTTCGCAAGTCTATTTTAAATAGCTCTTTTCAATCTCAATTTTTCGCAC 2881
Qy 2719 GGGCATCTCGAGATCGGTGAGCGGGGTGAGCATCGAGAGTGTGTGAGGAACGAGCA 2778
Db 2882 TGTATTTCTGAANAATGAGGTGGAGTGAGCATTTGAAGAAATGGTGGCGAATGAACA 2941
Qy 2779 GTTCTGGGTCAATCGGGGCGTGTCCGCGCATCTCTTCGCCGTGCTGAGGGCCTGCTCAA 2838
Db 2942 GTTCTGGGTATTTGGAGGGGTTTCTGCACATTTTTCGAGTTTATTCAGAGTCTGCTCAA 3001
Qy 2839 GGTCTTCGCGGATCGACACCACTTCCCGTCACTCCAAAGGCCACCGCGACGAGGA 2898
Db 3002 GGTACTGGCAGGCATTTGATACAAATTTCAAGTCACTGCGCAAGGCATC-----AGATGA 3055
Qy 2899 CGACGAGTTCGCGAGCTCTACGCTTCAAGTGGACCACTGCTCTCATCTCCGCCACAC 2958
Db 3056 CGGTGAGTTTGGGGAACCTGTATGATTCAAATGAACCACTCTCTCATTTCTCTCTACAC 3115

Qy	2959	GCTGCTCATATTAAAGTCATCGCGCTCGTGGCCGGCATCTCCGACGCCCATCAACAACGG	3018
Db	3116	CCTGCTTGTTCATCAACCTTGTGGGGTGGTTGTTGGCGTAGCAGATGCAATCAACAATGG	3175
Qy	3019	GTACCAAGTCCTGGGGGCCCTCTTCGGCAAGCTCTTTCGCCCTTCTGGGTCAATCGTCCA	3078
Db	3176	ATTTCAGTCATGGGGTCTCTCTTGGGTAAAGCTTTTCTTTGCATTTCTGGGTCAATTGTGCA	3235
Qy	3079	CCTCTACCGTTTCCTCAAGGGGCTCATGGGGCCGAGAACAGGACGCCACCGTTCTTGT	3138
Db	3236	CCTGTATCTTTCCTCAAGGCTCTCATGGGCGAGAACCCGACCATCGTGGT	3295
Qy	3139	CATCTGGTCCATTCTGTCCTCCATCTTCTCCCTGCTCTGGGTCAAGATCGACCCCTTT	3198
Db	3296	TATTTGGTCAATTCTCTGTCATCTGTTTCTCTCTTTCTGGGTAAGAATTGATCCTTT	3355
Qy	3199	CATCGTCAGGACCAAGGGCCCGACGTCAGGCAAGTGTGGCATCAATTGCTGAGCTGTTTA	3258
Db	3356	CTTGAGTAAGGTTAAAGGCCGAGATACTAAACAATGGGCATCAACTGCTGATTTCTTTG	3415
Qy	3259	TTA 3261	
Db	3416	ATA 3418	

Search completed: March 11, 2005, 21:58:40
Job time : 9704 secs

QY 2332 CACGAGGACATCTGACGGGTTCAAGATGCACTGCGCGGGTGGCGCTCCGTTGTA CTG 2391
| | | | |
Db 2279 GACGAGGACATCTTAACGGGTTCAAGATGCACTGCGCGGGTGGCGCTCCGTTGTA CTG 2338
| | | | |
QY 2392 CATGCCAAGCGCGCGGTTCAAGGGTTCGCGCCGATCAATCTATCGGACCGTCTCAA 2451
| | | | |
Db 2339 CACGCCGCGAGGGCGGCAATCAAGGGTTCGCGCCGATCAATCTGTCGATCGTCTGCA 2398
| | | | |
QY 2452 CCAAGTCTCGGTGGGCGGTGGGTCGCGATCTTTCAGCGCGGACAGCCCTT 2511
| | | | |
Db 2399 CCAAGTCTCGGTGGGCGGTGGGTCGCGATCTTTCAGCGCGGACAGCCCTT 2458
| | | | |
QY 2512 GCTGTACGGTACAAAGACGCGACCACTCAAGTGGCTGAGCGCTTCGCTACATCAAC 2571
| | | | |
Db 2459 CTGGTACCC---CTATGGCGCGCCCTCAAGTGGCTGAGCGCTTCGCTACACCAAC 2514
| | | | |
QY 2572 CACCATCTACCCCTTCACTCGCTCCGCTGCTCGCTACTGCAACCTCCCGCGGCTG 2631
| | | | |
Db 2515 CATGCTTACCCCTTCACTCGCTCCGCTGCTCGCTACTGCAACCTCCCGCGGCTG 2574
| | | | |
QY 2632 CTTCTTCAACGGCAAGTTCATATGCGCTCGATTCGATTCAGCGCTTCGCGAGCTCTTTCAT 2691
| | | | |
Db 2575 CTTCTTCAACGGCAAGTTCATATGCGCTCGATTCGATTCAGCGCTTCGCGAGCTCTTTCAT 2634
| | | | |
QY 2692 CGCCCTTTCATGTCATCTTCGCGACGGGATCTCGAGATCGGTGGAGCGGGTGA 2751
| | | | |
Db 2635 AGCGCTTTTCTGTCGATCATCGCGACGGGGTCTCGAGCTCGGTGGAGCGGGTGA 2694
| | | | |
QY 2752 CATCGAGGAGTGTGGAGGAACGAGCAGTCTTGGGTCTATCGGGCGGTGTCGCGCATCT 2811
| | | | |
Db 2695 CATCGAGGAGTGTGGAGGAACGAGCAGTCTTGGGTCTATCGGGCGGTGTCGCGCATCT 2754
| | | | |
QY 2812 CTTTCGCGTCTGCGAGGCGCTTCAAGTCTCTCGCGGGATCGACACCACTTCAACCGT 2871
| | | | |
Db 2755 GTTCGCGTCTTCAAGGCGCTCTCAAGTGTCTCGGGCGGTGTCGACCACTTCAACGGT 2814
| | | | |
QY 2872 CACTTCAAGGACACCGCGACGAGGA- CACAGATTCGCGAGCTCTACGCTTCAAGT 2930
| | | | |
Db 2815 GACGTCACAAAGCCCGCGCGACGAGACCGCGGTTCGCGAGCTCTAACTGTTCAGT 2874
| | | | |
QY 2931 GGACACGCTCTCATCCGCGCCACACGCTGTCTCATTAAGTCTATCGGGTCTGTGG 2990
| | | | |
Db 2875 GGACGAGCTGTGTGTCGCGCGAGCAGCTGATCATCAATGTTGGGATCTGTCG 2934
| | | | |
QY 2991 CCGGATCTCCGAGCGCATCAACACGGGTACAGTCTCTGGGGCCCTCTTCGGCAAGC 3050
| | | | |
Db 2935 CCGGGGTGTCGAGCGCGGTGAACAACGGGTACGGGTCTGGGGCCGCTGTTCGGGAAGC 2994
| | | | |
QY 3051 TCTTCTTCGCTTCTGCGTATGTCACCTTACCGGTCTCTCAAGGGGCTCATGGGGC 3110
| | | | |
Db 2995 TCTTCTTCTCTTCTGCGTATCTCCACCTTCTACCCCTTCTCAAGGGGCTCATGGGGA 3054
| | | | |
QY 3111 GCCAGAACGAGCGCCACCGTGTGTCATCTGCTCCATTCGCTGGCTCCATCTTCT 3170
| | | | |
Db 3055 GGACAGAACCGAGCGCCACAAATGTGCTGTCTGCTCAACCTCTCGCTCCATCTTCT 3114
| | | | |
QY 3171 CCTGTCTCTGGGTGAGGATCGACCTTTTCATGCTCAGGACCAAGGGGCCGAGCTCAGGC 3230
| | | | |
Db 3115 CCTGTCTCTGGGTGAGGATCGACCTTTTCATGCTCAGGACCAAGGGGCCGAGCTCAGGC 3174
| | | | |
QY 3231 AGTGTGACATCAATGTGAGCTGTT 3256
| | | | |
Db 3175 CATGCGGGGTCTGCTGTGAGCTGCT 3200
| | | | |

RESULT 3

AAZ99512

ID AAZ99512 standard; DNA; 3746 BP.

XX

AC AAZ99512;

XX

DT 03-JUL-2000 (first entry)

XX

DE DNA encoding a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
OS
XX Zea mays.
FH
FT Key Location/Qualifiers
FT CDS 321..3548
FT /*tag= a
FT /product= "cellulose synthase"
FT /transl_except= (pos: 1800..1802, aa: Xaa)
FT /note= "no termination codon given; Xaa is an unspecified
FT amino acid"
XX
PN WO200009706-A2.
XX
XX 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US018760.
XX
PR 17-AUG-1998; 98US-0096822P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX P-PSDB; AAY84114.
XX WPI; 2000-224343/19.
DR
XX
XX New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants and
PT to produce transgenic plants expressing the novel protein.
XX
XX Claim 1; Page 137-141; 119pp; English.
XX
XX The present sequence encodes a maize cellulose synthase polypeptide. The
CC cellulose synthase can be used for the improvement of stalk quality for
CC improved stand or silage. It also provides an increased concentration of
CC cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants and
CC seeds expressing the cellulose synthase. The polynucleotide is used for
CC modulating, preferably increasing, the level of the synthase in a plant
CC cell. The plants are preferably monocots. The polynucleotide is also used
CC as a probe or primer in the detection quantitation or isolation of gene
CC transcripts. The probes are useful in detecting deficiencies in the level
CC of mRNA in screenings for desired transgenic plant, for detecting or
CC mutations in the gene, for monitoring upregulation of expression or
CC changes in enzyme activity in screening assays of compounds, for
CC detection of any number of allelic variants of the gene, or for use as
CC molecular markers in plant breeding programs. The isolated nucleic acids
CC of the present invention can also be used for recombinant expression of
CC their encoded polypeptides or for use as immunogens in the preparation
CC and/or screening of antibodies. The proteins can be employed in assays
CC for enzyme agonists or antagonists of enzyme function or for use of
CC immunogens or antigens to obtain antibodies specifically immunoreactive
CC with a protein
XX
XX Sequence 3746 BP; 861 A; 976 C; 1006 G; 902 T; 0 U; 1 Other;

Query Match 40.6%; Score 1398.8; DB 3; Length 3746;
Best Local Similarity 67.2%; Pred. No. 3.8e-222;
Matches 2185; Conservative 1; Mismatches 923; Indels 143; Gaps 9;

QY 133 GAACGAGCTGCTGCTGATCGGGGCCACGAGGACCCACGCGCTCGGGGGCTGACGG 192
| | | | |
Db 308 GAGCTCGCTTGCATGAGGGGCGACGCGACGCGTGAAGTCGGGAGGCGCGGTGCGG 367
| | | | |
QY 193 GCAGGTGTGCGAGATATCGGCGCAGGAGTTCGGGCTCACCGTGCACCGCGACCTTCTCGT 252
| | | | |
Db 368 ACAGGTGTGCGAGATCTGCGCGCAGCGCGTGGGACACACGCGGAGGGGACGCTTCGC 427
| | | | |
QY 253 CGCCTGCAACGAGTGTGCGGCTTCCCCTGTGTGCCGCGCTTCTGTACGAGTACGAGCGCGGGA 312
| | | | |

Qy	2337	AGGACATCTCGAGCGGGTTCAAGATGCACTGCGCGGGTGGCGTCTCGTGACTGATCG	2399
Db	2647	AAGACATTTCTCACCGGATTCAGATGACACGCGGAGCTGGCGGTGCGATCTACTGCAATCG	2706
Qy	2397	CGAAGCGGGCGGGTTCAGGGGTGCGCGCGATCAATCTATCGAGCGGTCTCAACACGG	2456
Db	2707	CCAAGCGGCCAGCTTTTCAGGGGTCTGCCCCCATCAATCTTTTGGAGCGTCTGAACACGG	2766
Qy	2457	TGCTCCGGTGGGCGCTGGGGTTCGGTTCGAGATCTTTCTTCAGCGCGCACAGCCCCCTCGTGT	2516
Db	2767	TGCTCCGGTGGGCTCTTGGGTTCGGTGGAGATCTCTTCACGCGGCATCTGCCCTCTGTGT	2826
Qy	2517	ACGGCTACAGAAACGGCAACCTCAAGTGGCTGGAGCGGTTGCGCTATCAACAACACCA	2576
Db	2827	ACGGCTAC--GGAGGGCGGGCTCAAGTTCTTGGAGAGATTCGCGTATCAACAACACCA	2883
Qy	2577	TCTACCCCTTCACCTCGCTCCGCTGCTCGCTACTCGACCTCCCGCGCTGCGCTCC	2636
Db	2884	TCTACCCGCTCACGCTCAATCCGCTTCTCATCTACTGATCTCTGCCCGCATCTGTCTGC	2943
Qy	2637	TCACCGGCAAGTTCAATGATCGCGTTCGATAGACAGTTTCGCGAGCCTCTTTTCATTCGCCC	2696
Db	2944	TCACCGGAAGTTCAATCATTCAGAGATCAGCACTTCGCGAGCATCTGGTTTCATCTCCC	3003
Qy	2697	TCTTATGTGTCATCTTCGAGACGGGCATCTCTGGAGATGCGGTGAGCGGGTGAACATCG	2756
Db	3004	TCTTTCATCTCGATCTTGCCACGGGCATCTCTGGAGATGAGGTGAGCGGGGTGGGCATCG	3063
Qy	2757	AGGAGTCGTGGAGAAACGACAGTTCTGGGTTCATCGCGGGGTGTCGCGCATCTCTTCG	2816
Db	3064	ACGAGTGGTGGAGAAACGACAGTTCTGGGTTCATCGGGGTGATCTCGCGCACCTCTTCG	3123
Qy	2817	CCGTCTGTGAGGGCTGTCTAAGGTCTCTCGCGGGATCGACACAACTTCACCGTCACTT	2876
Db	3124	CCGTGTTCCAGGGCTGTCTAAGGTGTGGCGGCATCGACACAACTTCACCGTCACTT	3183
Qy	2877	CAAAGGCCACGGCGACGAGAGACGAGTTTCGCGAGCTCTACGCTCTCAAGTGGACCA	2936
Db	3184	CAAAGGCCTCG---GACGAGACGGCACCTTCGCGGAGCTGTATCATGTTCAAAGTGGACGA	3240
Qy	2937	CGCTCTCATCCCGCCACACACGCTGTCTCATATTAACGTATCATCGGGGTGTTGGCCGCGCA	2996
Db	3241	CGCTCTTGATCCCCGCCACCACTCTTGATCATCAACTGTTGTTGGGTGTTGGCCGCGCA	3300
Qy	2997	TCTCCGACGCCATCAACAAACGGGTACCAAGTCTTGGGGGCCCTCTTCGGCAAGCTCTTCT	3056
Db	3301	TCTCTACGCCATCAACAGCGGATACCAAGTGTGGGGCCGCTCTTCGGCAAGCTCTTCT	3360
Qy	3057	TCGCGTCTTGGGTCAATGTGTCACCTCTACCGGTTCTCAAGGGGCTCATGGGGGCGCAGA	3116
Db	3361	TCGCGTCTTGGGTCAATGTGTCACCTGTACCGGTTCTCAAGGGGCTCATGGGGGCGCAGA	3420
Qy	3117	ACAGGAGCCACACGGTTGTGTATCTGGTTCATCTGCTGGCTCCATCTTCTCCCTCGC	3176
Db	3421	ACCGACCCCGACCATCGTGTGTGTGGGCGATCTGCTGGGCTCCATCTTCTTCTCTTCG	3480
Qy	3177	TCTGGGTGAGGATCGACCTTTTCATCTGTCAGGAACAAAGGGGCCCGGAGCGTCAGGCGTGTG	3236
Db	3481	TGTGGGTTCGATCGACCCCTTCACACCGCGTCACTGGGCCCGGATACCGACGAGTGTG	3540
Qy	3237	GCATCAATTTGCT	3248
Db	3541	GCATCAACTGCT	3552

XX	DNA encoding a maize cellulose synthase.
DE	
XX	
KW	Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW	transgenic plant; plant breeding marker; ss.
XX	
XX	Zea mays.
OS	
XX	
XX	Location/Qualifiers
FT	321..3449
FT	/*tag= a
FT	/product= "cellulose synthase"
FT	/transl_except= (pos: 1800..1802, aa: Xaa)
FT	/note= "no termination codon given; Xaa is an unspecified
FT	amino acid"
XX	
XX	WO200009706-A2.
PN	
XX	
XX	24-FEB-2000.
PD	
XX	
XX	16-AUG-1999; 99WO-US018760.
PF	
XX	
XX	17-AUG-1998; 98US-0096822P.
PR	
XX	
XX	(PION-) PIONEER HI-BRED INT INC.
PA	
XX	
XX	Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
PI	
XX	
XX	WPI; 2000-224343/19.
DR	
XX	P-PSDR; AAY84119.
DR	
XX	
XX	New genes which encode maize cellulose synthase polypeptides in plants
PT	useful for modulating the expression of cellulose synthase in plants and
PT	to produce transgenic plants expressing the novel protein.
XX	
XX	Claim 1; Page 176-181; 119pp; English.
PS	

RESULT 4
AAZ99527
ID AAZ99527 standard; DNA; 3746 BP.
XX
AC AAZ99527;
XX
DT 03-JUL-2000 (first entry)

Qy	253	CGCCTGCACGAGTGGCGCTTCCCGCTGTGTCGGCCCTGTCTTACGAGTACGAGCGCCCGGGA	312
Db	428	CGCCTGGGACGCTCTGCGGGTTTCGGGTGTGCGGCCCTGTCTACGAGTACGAGCGCAAGGA	487
Qy	313	GGGCACGCAAACTGCCCCAGTGCAGACGCGCTACAAGGCCTCAAGGGGAGCCCGAG	372
Db	488	CGGCACGAGCGGTGCCCCAGTGCAGACCAAGTACAAGCGCCACAAGGGGAGCCCGCGC	547
Qy	373	GGTTGCGGGGACGATACGAGGAGGACATCGACGACTGGAGCACGAGTTCAACATCGA	432
Db	548	GATCCGTGGGAGGAGGAGACGACCTGATGCCGATAGCGATTTCAATTACCTTGCATC	607
Qy	433	CGACGAGAATCAGCAGAGGCGAGCTGGAGGGCAACATGCAGAAACGACGATCA-----	485
Db	608	TGGCAATGAGGACCAGAAGCAGAAGATTGCCGACAGAAATGCGCAGCTGGCGCATGAACGT	667
Qy	486	-----CCGAGGCCATGCTGCACGGCAGGATGAGCTACGGGAGGGGGCCCCG-----	530
Db	668	TGGGGGAGCGGGGATGTTGGTCCGCCCAAGTATGACAGTGGCGAGATCGGGCTTACCAA	727
Qy	531	--ACGACGGCAGCGGCAACAACCCCGCAGATCCCGCCCATCATCAACCGGCTCCCGGCTC	588
Db	728	GTATGACAGTGGCGAGATTCTCGGGGATACATCCCATCAGTCACTAAACAGCCAGATCTC	787
Qy	589	CGTGCCGGTGAGCGGTGAGTTTC--GATTACCAACGGGTATGGCCACGGCGAGGTCTCGT	647
Db	788	AGGAGAAATCCCTGGTGTCTCCCTGACCACATCATGATGTCCCCCACTGGGAACATTTG	847
Qy	648	CTTCCCTGCACAGCGCATCCATCCGTACCTGTGTCTGTAGCCAGGGAGTGCACAGTGGG	707
Db	848	CAACGCTGCTCCATTTC--CTATGTGAACCAATTGCCCAAAATCCGTCAAGGGAGTTCTCTG	906
Qy	708	ACGAGAAAGAAGTAGTGAGCTGGGAAGGAGAGGATGGAGCAGCTGGAACTCCAAGCAGGCA	767
Db	907	GTACCATTTGGGAATGTTGCCCTGGAAAGAGAGGTTGATGGCTGGAAATGAACGAGACA	966
Qy	768	-----TCC	770
Db	967	AGGGGACGATTTCCATGACGAATGGCACAGCATTGTCTCTCTGAGGTCGGGGTGTG	1026
Qy	771	TCGCGGGGGCCGATCCCGAGACATGAGCGCGAGCTGGGCNCTGAAACGACGAGGCGGA	830
Db	1027	GTGATATTGATGCAATCAACTGATTACAACATGGAAGATGCTTATTGAACGACGAAACT	1086
Qy	831	GGCAGCCGCTGTCCGAGGAAGTGTGCGATCGGCTCGACAAAGGTGAAACCCGTACCGGATG	890
Db	1087	GACAGCTCTATCTAGAGAAATTCCATTCTCTCCAGGATTAATCCATCAGGATGG	1146
Qy	891	TGATCGTGTGCTCTGTTGTGCTCGCTTCTTCTCCGATACCGGTATCCTGCACCCCG	950
Db	1147	TCATTGTGCTGCGATTGATTGTTCTAAGCATCTTCTTGGCATACCGTATCACAAATCCTG	1206
Qy	951	TCCCGGAGCCATCGGGCTGTGGCTGTCTCCATCATCTCGGAGATCTGGTTGCGCATCT	1010
Db	1207	TGCGCAATGCAATACCCATTATGGCTTCTATCTGTATTATGTGAGATCTGGTTGCTCTTT	1266
Qy	1011	CCTGGATCTCGACAGTTCCCAAGTGGTTCCCATCGACCGCGAGACCTACTCGACCC	1070
Db	1267	CGTGGATTTGGATTCAGTTCCCTAAGTGGTTTCCAATCAACCGGGAGACGTACTTGGATA	1326
Qy	1071	GCCTCTCCCTCAGGTACGAGGGAAGGGAGCGCTCGCTGCTGTCGGCGGTGGACCTGT	1130
Db	1327	GGCTGGCATTTAAGTATGACCGGAAGGTGAGCCATCTCAGTTGGCTGCTGTTGACATTT	1386
Qy	1131	TCGTGAGACGGTGGAACCGGCTCAAGGAGCGCGCTGGTGAACGCGCAACACCGTGTCT	1190
Db	1387	TCGTGAGTACAGTCGACCAATGAAGGAGCTCTCTTGTCACTGCGCAATACCGTGTCTAT	1446
Qy	1191	CCATCTCGCGTAGATACCCGTGACAGGTTCTCTCTGATGCTTCGACGAGCGCG	1250
Db	1447	CCATCTTGTGTGGATTACCCCTGTGATGAAGGTCTCTTGTCTATGTATCTGATGATGGAG	1506
Qy	1251	CGTCGATGCTGACGTTTCGAGTTCGCTGTGGAGACGGCGGAGTTTCGCGCGAAGTGGGTGC	1310

Qy	253	GCCTGTGCAA	CGAGTGTGCGGCTTCCCGCTGTGTCGGGCCCTGTCTACGAGTACGAGCGCCGGGA	312
Db	445	CGCCTGGGAC	CGTCTGTGCGGGTTTCGGGTGTGCGGCCCTGTCTACGAGTACGAGCGCAAGGA	504
Qy	313	GGGCACCGA	CACTGCCCCAGTGCAGACGCGTCTACAAGGCTCAAGGGGAGCCCGAG	372
Db	505	CGGCACGAG	CGGTGCCCCAGTGCAGACCAAGTACAAGCGCCACAAGGGGAGCCCGGC	564
Qy	373	GGTTGCGGG	ACGATGACGAGGAGGACATCGACGACCTCGGAGCAGCAGTTCACATCGA	432
Db	565	GATCCGTGG	GAGGAGGACGACACTGATGCCGATAGCGACTTCAATTACCTTGCATC	624
Qy	433	CGACGAA	TACAGCAGGAGCGTGTGGAGGGCAACATCGAAGCAGCCAGATCA-----	485
Db	625	TGGCAATC	AGACCAAGCAGAGGATGCGGATGCGGATGCGGCGCATGAACGT	684
Qy	486	-----	CCGAGGCGATCTGCACGGCAGGATGAGCTACGGAGGGGGCCCCG-----	530
Db	685	TGGGGG	CAGCGGGATGTTGGTGCCTCCCAAGTATGACAGTGGCGAGATCGGGCTTACCAA	744
Qy	531	--	ACGACGGGACGGCAACAACCCCGCAGATCCCGCCCATCATCACCGGCTCCCGCTC	588
Db	745	GTATGAC	GATGGCGAGATTCTCGGGATACATCCCATCAGTCACTAACAGCCAGATCTC	804
Qy	589	CGTGCCGGT	GAGCGGTGAGTTTCG-GATTACCAACGGGTATGGCCAGCGGAGGTCTCGT	647
Db	805	AGGAGAA	TCCCTGCTCCCTGACCATCATATGATGCCCACTGGGAACATTGG	864
Qy	648	CTTCCCTGC	ACAACGGCATCCATCCGTACCTGTGTCTGACGCCAGGAGTGCAGTGGG	707
Db	865	CAAGCGT	GTCTCAATTTC-CCTATGTGAACCATTCGCCAAATCCCGTCAAGGGAGTCTCTG	923
Qy	708	ACGAGAA	GAAGAAGTGCAGCTGGAAGGAGAGGATGAGACGACTGGAAGTCCAAAGCAGGCA	767
Db	924	GTAGCAT	TGGGATGTGCTTGGAAAGAGAGGGTTGATGGCTGGAAATGAAGCAGGACA	983
Qy	768	-----	-----TCC	770
Db	984	AGGGACG	ATTCCCATGACGAATGGCACAGCAATTGCTCCTCTGAGGGTCGGGGTTG	1043
Qy	771	TCGGCGGG	CGCCGATCCCGAGACATGAGCGCGAGTGGCACTGAACGACGAGGCGA	830
Db	1044	GTGATATT	GATATCAACTGATTACAACTATGGAAGATGCTTATTGAACGACGAAATC	1103
Qy	831	GGCAGCG	CTGTTCGAGGAAGTGTGATTCGCTCGACAAAGGTGAACCCGTACCGATGG	890
Db	1104	GACAGCT	CTACTATAGAAAGTTCCACTTCCTCCTCCAGATAAATCCATACAGATGG	1163
Qy	891	TGATCG	TGGTCTGTGTGCTCGCTTCTTCCTCCGGTACCGGTATCCTGCACCCCG	950
Db	1164	TCATTGT	GTGCGATTGTTTCTAAGCATCTTCTTGCACTACCGGTATCACAATCCCTG	1223
Qy	951	TCCCGG	AGCCATCGGCTGTCTCATCTCTGATCTCGGAGATCTGGTTCGCCATCT	1010
Db	1224	TGCGCA	ATGCAATACCCATTTAGGCTTCTATCTGTATTATGTGAGATCTGGTTGCTCTT	1283
Qy	1011	CTTGGAT	CTCTGACAGTTCCTCAAGTGGTTCCCATCGACCGGAGACGTACTCGACC	1070
Db	1284	CGTGGAT	TATGGATCAGTTCCTCAAGTGGTTTCCAAATCAACCGGGAGACGTACTTGATA	1343
Qy	1071	GCCTCT	CCCTCAGTACGAGGGAAGGGAGCGCTCGCTGTCTGGCGGTGGACCTGT	1130
Db	1344	GGCTGG	CAATTAAGGTATGACCGGGAAGGTGAGCCATCTCAGTTGGCTGTCTTGACATTT	1403
Qy	1131	TCGTGAC	ACGCTGGATCCCGCTCAAGGAGCGCGCTGGTGAACCGCCAAACACCGTGTCT	1190
Db	1404	TCGTCA	TAGTCAAGTCCCAATGAAGAGCGCTCTCTTGTCACTGTCCTCAATACCGTGTCT	1463
Qy	1191	CCATCT	TCGGGTAGATACCCGTGACAGGTCTCTCTGATAGTCTCGACGAGCGCG	1250
Db	1464	CCATTCT	TGCTGGATTACCTGTGATGAAGGTCTCTTGTCTATGTATCTGATGATGGAG	1523

Db 2604 GCTATAGGCAAGACTGAATGGGAACTGAGATCGGCTGATCTACGGTTCTGTGACG 2663
Qy 2337 AGGACATCTCTGACGGGGTTCAAGATGCACTCCGCGGGTGGCGTCCGTGTACTGCAATGC 2396
Db 2664 AAGACATTTCTACCCGGATTCAAGATGACGCGCGAGGCTGGCGGTGATCTACTGCAATGC 2723
Qy 2397 CGAAGCGGGCGGGTTCAAGGGGTGGCGCGCGATCAATCTATCGGACCGTCTCAACCAAG 2456
Db 2724 CCAAGCGGCCAGCTTTCAAGGGGTCTGCCCCCATCAATCTTTTCGAGCCGTCTGAACCAAG 2783
Qy 2457 TGCTCCGGTGGCGCTGGGGTCCGTTCGAGATCTTCTTCAGCGCGCAGACCCCTGCTGT 2516
Db 2784 TGCTCCGGTGGCTCTTGGGTCCGTGGAGATCTCTTCAGCGCGCAGTGGCCCCCTGTGGT 2843
Qy 2517 ACGCTACAAGAACCGCAACCTCAAGTGGCTGGAGCGCTTCGCCCTATCATCAACACCA 2576
Db 2844 ACGGCTAC--GGAGGGCGGCTCAAGTTCTTGGAGAGATTCGCGTATCATCAACACCA 2900
Qy 2577 TCTACCGCTTCACTCGCTCCGCTGCTGGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCT 2636
Db 2901 TCTACCGCTTCACTCGCTCCGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTGCTGCT 2960
Qy 2637 TCACCGCGCAAGTTCATCATGCGCTCGATTAGCACGTTGCGCAGCTCTTCTTCATCGCC 2696
Db 2961 TCACCGGAAGTTTCATCATTCAGAGATCAGCAACTTCGCCAGCATCTGTTTCATCTCCC 3020
Qy 2697 TCTTCATGCTCATCTTCGCGACGGGCTCTTGGAGATCGGTTGAGCGGGTGGAGCATCG 2756
Db 3021 TCTTCATGCTCATCTTCGCGACGGGCTCTTGGAGATGAGTGGAGCGGGTGGAGCATCG 3080
Qy 2757 AGGAGTGTGAGGAAAGAGCAGTTCTGGGTTCATCGCGCGGTGTCGCGCATCTCTTCG 2816
Db 3081 ACAGTGTGTGAGGAAAGAGCAGTTCTGGGTTCATCGCGCGGTGTCGCGCATCTCTTCG 3140
Qy 2817 CCGTGTGCGAGGGCTGCTCAAGTCTCTCGCGGGATCGACCACTTACCGTCACT 2876
Db 3141 CCGTGTGCGAGGGCTGCTCAAGTCTCTCGCGGGATCGACCACTTACCGTCACT 3200
Qy 2877 CCAAGGCGCACCGCGACGAGGACGAGTTTCGCGGAGCTCTAGCGCTTCAAGTGGACCA 2936
Db 3201 CCAAGGCGCTCG--GACGAGGACGGGACTTCGCGGAGCTGTACATGTTCAAGTGGACGA 3257
Qy 2937 CGTCTCTCATCCGCGCACCAAGCTGCTCATTAAGTCAATCGGCGTGTGCGCGGCA 2996
Db 3258 CGTCTCTCATCCGCGCACCAAGCTGCTCATTAAGTCAATCGGCGTGTGCGCGGCA 3317
Qy 2997 TCTCGAGCGCATCAACAGGCTACAGTCTCTGGGGGCGCTCTTTCGCGAGCTCTTCT 3056
Db 3318 TCTCTTACGCGCATCAACAGGCTACAGTCTCTGGGGGCGCTCTTTCGCGAGCTCTTCT 3377
Qy 3057 TCGCTTCTTGGGTTCATGTCACCTCTACCGGTTCTCAAGGGCTCATGGGGGCGCAGA 3116
Db 3378 TCGCTTCTTGGGTTCATGTCACCTCTACCGGTTCTCAAGGGCTCATGGGGGCGCAGA 3437
Qy 3117 ACAGAGCGCCACCGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3176
Db 3438 ACCGACCGCCACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3497
Qy 3177 TCTGGGTGAGATCGACCTTTCATGCTCAGGACGAGGGGCGCGGACGTCAGGCAAGTGTG 3236
Db 3498 TGTGGGTTCGATCGACCCCTTTCACCCCGCTCACTGGGCGCGGATACCCAGACGTTGTG 3557
Qy 3237 GCATCAATTGCT 3248
Db 3558 GCATCAACTGCT 3569

RESULT 6

AAA67114

ID AAA67114 standard; DNA; 3851 BP.

XX

AC AAA67114;

XX
DT
XX
DE
XX
KW
KW
XX
OS
XX
PN
XX
PD
XX
PF
XX
PR
PR
XX
PA
XX
PI
XX
DR
DR
XX
PT
PT
PT
XX
PS
XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
SQ

31-OCT-2000 (first entry)

Pinus radiata cellulose synthase nucleotide sequence SEQ ID NO:69.
Eucalyptus grandis; pinus radiata; Monterey pine; modification;
plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;
transgenic plant; ds.
Pinus radiata.

WO200022092-A2.

20-APR-2000.

08-OCT-1999; 99WO-NZ000169.

13-OCT-1998; 98US-00170862.

11-AUG-1999; 99US-0148426P.

(GENE-) GENESIS RES & DEV CORP LTD.
(FLET-) FLETCHER CHALLENGE FORESTS LTD.

Bloksberg LN;

WPI; 2000-339328/29.

P-PSDB; AAB16307.

New genes encoding proteins involved in a plant polysaccharide
biosynthetic pathway, useful for modulating or altering the
polysaccharide content, composition or structure of the plant.

Claim 1; Page 71-72; 301pp; English.

The present invention describes isolated polynucleotides (PN) comprising
a sequence selected from one of 835 nucleotide sequences given in
AA67073 to AA67907, their (reverse) complements, sequences producing an
Expectation (E) value of 0.01 or less compared to the 835 sequences,
sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the
835 sequences or sequences that are degenerately equivalent or allelic to
the 835 sequences. The polynucleotides are used to modify the activity of
a polypeptide involved in a polysaccharide biosynthetic pathway in the
plant. They are especially used to modulate or alter the polysaccharide
content, composition or structure of the plant. AAB16268 to AAB16340 are
proteins encoded by some of the polynucleotide sequence given in the
present invention

Sequence 3851 BP; 1040 A; 796 C; 944 G; 1071 T; 0 U; 0 Other;

Query Match

Best Local Similarity 39.7%; Score 1367; DB 3; Length 3851;

Matches 2197; Conservative 0; Mismatches 940; Indels 166; Gaps 10;

Qy 90 CGATGGAGGCGAGCGCGGCTGGTGGCTGCGACACCGGACGAGCTGGTGTGCA 149
Db 151 CAATGGAAGCCAGCGCGGCTTGTGCTGCTTCTATAACAGAAACGAGTGTGTGCA 210
Qy 150 TCCGGGCGCCACGAGGACCCCAAGCGCTGCGCGCTGAGCGGCGAGTGTGCGAGATAT 209
Db 211 TCCATGGACATGAGGAGCCGAGCCTTGAACAGTTGAGTGGCCAGCTCTGCCAGATT 270
Qy 210 GCGGCGACGAGGTCGGGCTCACGGTGGAGCGGACCTCTTCTGCTGCTCAACGAGTGG 269
Db 271 GTGCGGAGGACGTGGGCTTAAACACAGACGCGGAGCTGTTCGTGCTGCTGTAATGAGTGG 330
Qy 270 GCTTCCCGCTGTCGGGCTGCTACAGTACGAGCGCGGAGGCGACCGAGACTGCC 329
Db 331 GGTTCCTGCTGTCGGGCTGCTATGAGTACGAGAGACGAGAGGAAATCAGTCTGCGC 390
Qy 330 CCCAGTGCAGAGCGCGCTACAAGCGCTCAAGGGGAGCCCGAGGTTGCCGGGACGATG 389
Db 391 CGCAGTGCAATACTCGTTTACAGCGTCAAAAAGGAGTCCACCGGTTGGAAGGTGACGATG 450

QY 390 ACGAGGAGCATCGACGACCTGGAGCACAGGTTCAACATCTGACGACGAGAAATCAGAGA 449
Db 451 ATGAAGAAGACGTTGATGACATAGAAACATGAATTTTATGTGGAGACTTCAGCAAGAAGAAC 510
QY 450 GGACGCTGGAGGGCAACATGCAGAAACAGCAGATCACCGAGGCGATGCTGCACGCGAGGA 509
Db 511 GGCAG-----CAGATCACCGAGGCGATGCTCCACGAGCGCA 546
QY 510 TGAGCTACGGAGGGGCGGACGACGCGACGGCAACAAACCCCGCAGATCCCGCCCA 569
Db 547 TGAGCTATGCCGAGGTCGCGACGACGAAATTCGCAGATTGCTC--ATAATCCAGAGCT 604
QY 570 TCAATCACCGGCTCCCGTCCGTCG-----CGTGAGCGGTGAGTTCCGAT 615
Db 605 TCCTCCGCAATTCTCTGATCTTGCAAAACGGCCACTCGGTTGTGAGTGGGAGATTCCAAC 664
QY 616 TAC---CAACGGGTATGGCCACGCGAGGTCCTGCTCTTCCCTGCACAAGCGCATCCATCC 672
Db 665 GTCATACTACGCAGACMACCAATTCGTTGCAACCCCTGCAATGCTGAAGCGGTGTCATCC 724
QY 673 GTACCCCTGTCTTGAGCCAGGAGTGCCAAAGTGGGACGAGAGAAAG-----719
Db 725 AAGCTCCGACCGGGAGTGGAGGATCATATGGATCCAAACAGGATATTGGTTCTTA 784
QY 720 -----AAGTGAGCTGGAAGAGAGGATGGACGACTGGAAGTCCAAAGCAGGGCAT 768
Db 785 TGGCTTTGGGAAACGTGTCTTTGGAGGAGCGAGCGATGGTTGTATAAATCGAAGGAAACAA 844
QY 769 CCTCGGCGCGCGCGATCCGAG-----794
Db 845 ATCAGGCCAGTTGGATATGACGAGAGGAGATATCAATATATAATGCGGGGTTTCACCAAA 904
QY 795 -----ACATGGACCGCAGCTGGCACTGAAACGACGAGCGCAGGCGCGCT 840
Db 905 TGAGCCTGAAGATTATATTGATCCCGATATGCCAATGACCGATGAAGCAAGGCGACCT 964
QY 841 GTCAGGAAGGTGTCGATCCGCTCGACGAAGGTGAACCCGTACCGGATGTGATCGTGGT 900
Db 965 GTCCGAAAAAGTGCRAATTCCTTCAAGCAAAATAAATCCATACCGAATGGTCAATTGTAAT 1024
QY 901 GCGTCTGTTGTCTGCTGCTCTTCTTCGCTACCGTATCTCAGCCCGTCCGCGACGC 960
Db 1025 TCAGTCAATAGTCTGGGTATTTTCTCCGCTATCGTCTCCTGAATCCAGTGAAGATGC 1084
QY 961 CATCGGCTGTGGTCTGCTCCATCATCTGCGAGATCTGGTTCGCCATCTCTCTGGATCCT 1020
Db 1085 ATATGGGCTCTGGGCCACTTCTATCGTTTGTGAAATCTGTTGGCTTGTGATGATCT 1144
QY 1021 CGACCACTTCCCAAGTGGTTCGCCATCGACCGGAGACGTACTCGACCGCCTCTCCCT 1080
Db 1145 TGATCAAGTTTCCCAAGTGGTTCGCTATCATGCTGTGAAACGTATCCTGATCGACTGTCAAT 1204
QY 1081 CAGGTACGAGAGGAGGGAGCGCTCGCTGCTGTGGCGGTGACCTGTTGCTGTGAGCAC 1140
Db 1205 AAGGTACGAACGAGAAGCGAACATCAATGCTTGTGACCTGTTGACCTCTTTGTGAGTAC 1264
QY 1141 GGTGGACCCGCTCAAGGAGCGCGCTGTGTGACCGCAACACCGTGTCTTCCATCTCTCGC 1200
Db 1265 TGTAGATCCACTGAAGGAGCCTCCTTTGGTTACTGGCAATACAGTATTATCAATCCTTTC 1324
QY 1201 CGTAGACTACCCGTGACAAAGGTCTCTGCTAGTCTCGACGACGCGCGCTGATGCT 1260
Db 1325 AGTAGACTACCTGTAGACAATGTGCTCTGTTATGCTCTGATGACGAGCGCTGATGCT 1384
QY 1261 GAGCTTCGAGTCTGTGCGAGACGCGCGAGTTTCGCGCGAAGTGGGTGCCCTTCGCA 1320
Db 1385 TACTTTTGAATCTCTCTGAGACCTCAGAATTTGCCAANAATGGGTACCATTCGCA 1444
QY 1321 GAAGTTTCGGATCGAGCCCGCGCCCGAGTTCTACTTCTCGCTCAAGTTCGACTACCT 1380
Db 1445 GAAATTCGACATTTGAGCTCGCGCTCCGGAATCTATTCTCTCAGAAAAATTGACTATCT 1504
QY 1381 CAAGGACNAGGTGACGCCCACTTCTGTGAGGAGCGCGCCCATGAGAGAGATATGA 1440

Db 1505 GAAGGCAAAATTTCAACCCACCTTTGTCAAGAGCGCCGTGCCATGAAGAGAAATATGA 1564
QY 1441 CGAGTTCAAGGTCGGATCAACCGCTGTGTGGCCAAAGGCCATGAAGTGCCTCCGACAGGG 1500
Db 1565 AGAATTCAGAGTGGCATCAATCGGTGTGTGCAAGGCTCTAAAGTGCCCAAGGAAG 1624
QY 1501 GTGGATCATGAAGAGCGGACGCGCTGGCCCGGGAAACAACCCGCGACCAACCCCGCAT 1560
Db 1625 ATGGACAAATGCAAGACGGTACGCTTGGCTGTGTAATAATACCCGTGACCATCTCGTAT 1684
QY 1561 GATCAGGTGTTCTTGGGCCACAGCGCGGCCACGACCGAGGGCAACAGGTGCTCCCG 1620
Db 1685 GATCCAAGTGTCTTGGGTCACAGTGGCGGCTCGATACAGAAGGCAATGAGCTTCTCTG 1744
QY 1621 CTTCTGTGACGTCCTCCGCTGAGAAGCGCCGGGATTTCCAGCAACCAAGAAGSCCGCGC 1680
Db 1745 GCTAGTATATGTTTCTCTGTGAGAAGACCTGTTTTCAGCATCACAGAAGSCCGTGC 1804
QY 1681 CATGAACGCTCTGATTCGCGTCTCCGCGCTGTGTGACCAACGCGCCCATTCATGCTCAACTT 1740
Db 1805 CATGAATGCTTGTGGTTTCGGGTTTCTGCTGTCTCACCATGCTCCATTTATGCTGAATCT 1864
QY 1741 GGAATGTGATCACTACATCAACACAGAAAGGCATCCCGGAGGCCATGTGCTTCTCAT 1800
Db 1865 GGAATGTGATCACTACATTAACAATAGCAAGGCAATCAGGGAAGGCATGTGCTTTATGAT 1924
QY 1801 GGAACCTCAGTCTGGCGCGGAAGTCTGTACGTTTCAAGTTCCCGCAGAGGTTGACCGCAT 1860
Db 1925 GGAATCCTCAGTTTGGAGAAAGTCTGTATGTCTCAATTTCCCTCAGAGATTCGATGAT 1984
QY 1861 CGACGTGACGACCGCATACGCTAAACAGGAACACCGTCTTCTTCGACATCAACATGAAGGG 1920
Db 1985 TGATCGCAATGACCGTTACGCCAATCGAAACACCGTATTCTTTGATATCAACATGAAGG 2044
QY 1921 GCTGAGCGCATCCAAAGGCCCGGTGTAAGTGTGCGGACAGGGTGTGTTCCGGCGCCAGGC 1980
Db 2045 TCTGGATGGAATTCAGGGCCTGTATGTGGAACCTGGATGTCATGTTTCAGAGAACAAGC 2104
QY 1981 GCTCTACGGCTCAACCCCTCCAAAGGACCCAAAGAGGCCAAGATGTCGACCTGCGACATG 2040
Db 2105 TCTATATGGGTATGGGCTCCCAAGGCCCAAAACGTCCCAAGATGTTGATGATGTTG 2164
QY 2041 CTGCGCGTCTTTCGGCGCAAGAAAG-----2065
Db 2165 TCTCCCTTGTGCGTCTCTGTAAGAAAGTCTCCGAAGAAATAATAGTACGAAGAAAGTGC 2224
QY 2066 -CGGAACACGCCAAGGACGGGCTGCCGAGGCGCACCGCTGATATGGGAGTA-----GA 2118
Db 2225 AGGAATCCCAAGTCCCGCTACAAATCTGGACGGGATTCGAGGAAGGATGAAGGTTATGA 2284
QY 2119 TAGCGACAAAGGAGATGCTCATGTCCACATGAATTCGAGAAGCGGTTCCGGCAGTCCCG 2178
Db 2285 TGACGAAGAGCATGTTGTGATGAGCCACTAGACTTCGAGAAGAGTTTGGCCAGTCTTC 2344
QY 2179 GCGTTCGTCACTCGACGCTGATGGAGAAGCGCGCTCTCTCTCTGTCGAGCCCCCG 2238
Db 2345 AGCTTTTGTTCAAATCCACTCTGATGGAGAAATGGTGGTGTTCGCAAAACAGCAATCCAGC 2404
QY 2239 CGGCTCTCAAGGAGGCCATCCATGTCTCATGCTGCGCTACGAGGACGAAGCCCACTG 2298
Db 2405 TGAATTTGTTGAAGGAGGCTTATTCATGTCATGCTGTGATATGAAGACAAAACGGAATG 2464
QY 2299 GCGGCTGAGGCTGGGTGGATCTTACGGGTGATTCACGAGGACATCTCTGACGGGTTCAA 2358
Db 2465 GGGAAAAGAGCTTGGATGATCTATGGATCAGTCAAGAGGACATTCGACTGGATTCAA 2524
QY 2359 GATGCACTCCCGGCTGGCTCGGTGTAATGTCATGTCGAAAGCGGGCGGCTCAAGGG 2418
Db 2525 GATGCACACTCGAGGCTGGCGGTCCATTTACTGTATGCTGCCCCAAACGAGCAGCATTCAGAG 2584
QY 2419 GTGCGGCGCGATCAATCTATCGGACCGTCTCAACCAAGTGTCCGCTGGCGGCTGGGTC 2478

Db 511 GCGCGGCGACGAGGAGGACGGCGTTCGACGACCTGAGGAGCGAGTTGCGCTCGAGGA 570
Qy 436 CGAGAAATCAGCAGAGGCGAGCTGAGGGCAACATGCAAGAACAGCAGATCACCGAGGCGAT 495
Db 571 CGCGCGCGCCACGAGGACGACCGCAGTACGTGCGGAGTCCATGCTCAGGCGCGAGAT 630
Qy 496 GCTGCGCAGGAGATGAGCTACGGGAGGGGCCCGGACGACGGCGGACGGCAACAAACCC 555
Db 631 GAGCTACGGCGCGCGGCGGACGCGCACCCCGGCTTCAGCCCGCTCCCAACGTCGCGCT 690
Qy 556 GCAGATCCCGCCCATCATCACCGGCTCCCGCTCGGTGCGGTGAGGTGAGTTCCGAT 615
Db 691 CCTCACAAACGGCCAGATGTTGATGACATCCCGCGGAGCAGCAGCGCTCGTGCCTC 750
Qy 616 TACCAACGGGTATGCGCACGCGCAGGTCTCGTCTTCCCTGCACAAGCGCATCCATCCGTA 675
Db 751 CTACATGAGCGGCGCGCGCGGGGCAAGAGATCCACCCGCTCCCTTTCGCAGATCC 810
Qy 676 CCCTGTGTCGAGCCAGGGAGTGCCAAAGTGGGACGAGAAGAA----- 718
Db 811 CAACCTTCCAGTGCAACCGAGATCCATGGACCCGCTCCAAGGATCTGGCGCGCTACGGATA 870
Qy 719 -----GAAGTCAGCTGGNAGAGAGGATGGACGACTGGAGTCCAAGCAGGCACTCT--- 771
Db 871 TGGCAGCGTGGCTTGGAAAGAGAGAAATGGAGGGCTGGAAAGCAGAGAGCGCTGCA 930
Qy 772 -----CGGCGGCGCGCGCGGATCCCGAAGACATGGACCGCGACGTGGCACTGAA 819
Db 931 GCATGTCAGAGCGAGGCTGGCGGTGATTTGGGATGGCGACGATGCGCACTAAT 990
Qy 820 CGACGAGCGCAGGCGCGCTGTCGAGGAAAGGTGTCGATCGCTCGAGCAAGGTGAACCC 879
Db 991 GGAATGAAGCTAGGCAGCCATTTGTCAGAAAGTCCCTATATCATCAAGCCGAATTAATCC 1050
Qy 880 GTACCGGATGTCATCGTGTGCTGCTCGTGTGCTCGCTTCTTCTCGGTACCGTAT 939
Db 1051 CTACAGGATGATTCGTTATCCGGTTGGGTGTTTGGGTTCCTTTCACCTACCGAGT 1110
Qy 940 CTTGCAACCCCGTCCCGACGCCATCGGGCTGTGGCTCTCCATCATTCGCGAGATCTG 999
Db 1111 GATGCATCCGGCGAAGATGCAATTTGCAATTTGGCTCATATCTGTAATCTGTGAATCTG 1170
Qy 1000 GTTCGCCATCTCTGGAATCTCGACCAAGTTCCCAAGTGGTTCCCATCGACCGCGAGAC 1059
Db 1171 GTTTGGCATGTCCTGGAATTTGATCAGTTTCCCAAGTGGCTTCCAAATCGAGAGAGAC 1230
Qy 1060 GTACCTCGACCGCTCTCCCTCAGGTACGAGAGGAGGGAGCGCTCGCTGCTGCTCGC 1119
Db 1231 TTACCTGGACCGTTTGTCACTAAGGTTTGAAGAAGAGGTCAACCTCTCAGCTTGTCTC 1290
Qy 1120 GGTGGACCTGTCGTGAGCACGGTGGACCCGCTCAAGGAGCGCGCTGGTGACCGCCAA 1179
Db 1291 AATCGACTCTTTGTGATACGTTGATCCACAAAGGAACCTCCCTTGTTCACAGGAA 1350
Qy 1180 CACGCTGCTCTCATCTCCGCGTAGACTACCCGTTGGAACAAGTCTCTGCTACGTCCTC 1239
Db 1351 CACTGTCCTTTCCATCTCTGTGGATTAATCCGTTTGAAGAGTCTCTCTGCTATGTTTC 1410
Qy 1240 CGACGAGCGCGCTGAGTCTGAGTCTGAGTCTGCTGCTGCGAGACGCGCGAGTTCCGCG 1299
Db 1411 TGATGATGGTGTGCAATGCTTACGTTTGAAGCATTTGCTGGAACATCTGAAATTTGCAA 1470
Qy 1300 CAAGTGGGTGCCCTTCTGCAAGAAGTTTCGGCATCGAGCCCCCGCGCGAGTTCTACTT 1359
Db 1471 GAAATGGTTCCTTTCAGCAAAAAGTTTAATATCGAGCTCTGTCCTCTGAGTGGTACTT 1530
Qy 1360 CTCGCTCAAGGTGCACTACTCAAGGACAAAGGTGCGAGCCCACTTCGTGCGAGAGCGCG 1419
Db 1531 CCAACAGAGATAGACTACTGAAAGACAAGGTGCTGCTTCATTTGTTAGGAGAGGAG 1590
Qy 1420 CGCCATGAAGAGAGATATCAGGAGTTCAAGGTCCGAGTCAACCGCTGTCGCCAAGGC 1479
Db 1591 GCGATGAAGAGAGAAATACGAGGAATTCAAAGGTAAGGATCAATGCTTGGTTGCAAAAGC 1650

Qy 1480 CATGAAGGTGCGCGCAGAGGGGTGGATCATGAAGGACGCAACCGCTGGCCCGGGGAACAA 1539
Db 1651 CCAAAAGGTTCTCTGAGGAAGGATGACAATGCAAGATGGAAGCCCTGCGCTGGAACAA 1710
Qy 1540 CACCCGCGACCAACCCCGGCATGATCCAGGTGTTCTTGGGCCACAGCGCGGCCACACAC 1599
Db 1711 CGTACGCGATCATCTCTGGAATGATTCAGGTATTTCTTGGCCAAAGTGGCGGTGATGT 1770
Qy 1600 CGAGGCAACGAGCTGCCCCGCTCGTGTACGTCTCCGTGAGAACGCCCGGGATTCCA 1659
Db 1771 GGAAGAAATGAGTTGCTCGCTGCTGTTATGTCTCGAGAGAAAGAGGCCAGGTTATAA 1830
Qy 1660 GCACCAAGAAGAGCGCGCCCATGAACGCTCTGATTCGCGTCTCCCGCTGTCGACCAA 1719
Db 1831 CCATCACAAAGAGGCTGGTGCCATGAATGCACTGGTCCGTGCTCTGCTCTTATCAAA 1890
Qy 1720 CGCGCCATTCATGCTCACTTGGACTGTGATCACTACATCAACAAAGAGGCCATCCG 1779
Db 1891 TGTGTCATACCTATTGAACCTTGGACTGTGATCACTACATCAACAAATAGCAAGGCCATAA 1950
Qy 1780 GGAGGCCATGTCTTCTCATGACCCCTCAGGTCGSCCGGAAGGTCTGCTACGTTCAGTT 1839
Db 1951 AGAGGCTATGTGTTTCATGATGATCTTTGGTGGGAAAGTGTGCTATGTACAGTT 2010
Qy 1840 CCGCAGAGGTTTCGACGGCATCGAGTGCACGACCGATACGCTAACAGGAACACCGTCTT 1899
Db 2011 CCTCAGAGGTTTGATGGTATTGACAAAATGATCGATACGCTAACAGGAACGTTGCTCT 2070
Qy 1900 CTTTCGACATCAACATGAAGGGCTGACCGCATCCAGGCCCGGTGTAGCTCGGACAG 1959
Db 2071 TTTTGCATCAACATGAAGGTTTGGACGGTATTCAAGGACCCATTTATTTGGGTACTGG 2130
Qy 1960 GTGCGTGTTCGCGCGCAGCGCTCTACGCGCTACAAACCTCCCAAGGACCCCAAGAGGC 2019
Db 2131 ATGTGTTTTCAGACGGCAGGCACTGTATGTTATGATGCTCTTAAACGAGAGCCACC 2190
Qy 2020 -----CAAGATGGTGAACCTGCGACTGCTGCCGTGCTCGCGCGCAAGAGCG 2067
Db 2191 ATCAAGAACTTGCNACTGTGCGCCCAAGTGGTGCCTCTCTTGTGCTGCGCAGCAAGCAA 2250
Qy 2068 GAAACAGCCCAAGGACGGCTGCGGAG----- 2095
Db 2251 GAATAAAAGNAGACTACAAAACCAAGACGGAAGAAAGAAAGATTAATTTTCAAGAA 2310
Qy 2096 -----GGCACCGCTGATATGGAGT 2115
Db 2311 AGCAGAAACCCATCTCTGCATATGCTTTGGGTGAAATTTGATGAAGGTGCTCCAGGTGC 2370
Qy 2116 AGATAGCGCAAGAGAGATGCTCATGTCCCATGAACTTTGAGAAAGCGGTTCGGGCGATC 2175
Db 2371 TGATATCGAGAAAGCGCGAATCGTAAATCAACAGAAACTAGAGAGAAAATTTGGGCGATC 2430
Qy 2176 CGCGGGTTCGTCACGTCCAGCTGATGAGGAAGCGCGGGTCCCTCTCTGCTCGAGCCC 2235
Db 2431 TTTCTGTTTTGTGCGCATCAACACTTCTTGAGAACGGAGGACCTCGAAGCGCCAAGTCC 2490
Qy 2236 CGCGCGCTCTCAAGGAGGCCATCCATGTCTACGTGCTGCGGTACGAGGACAAAGACCGA 2295
Db 2491 AGCTTCTCTTGAAGGAAGCTATACATGTTATCAGTTCGGCTACGAGACAAAGACCGA 2550
Qy 2296 CTGGGGCTGAGAGCTGGGTGGATCTACGGGTGATCAACGGAGGACATCTGACGGGTT 2355
Db 2551 CTGGGGAAGAGATGGCTGGATTTACGGATCGATCACAGGATATCTTGACTGGAT 2610
Qy 2356 CAAGATGCACTCGCGGGTGGGCTCCGTGACTGTCATGCCGAAGCGCGCGGTCAA 2415
Db 2611 TAAGATGCACTGCCATGGCTGGCGGTCTATTTACTGTCATCCCGAAGCGGCTGCAATCAA 2670
Qy 2416 GGCGTCCGCGCGCATCAATCTTATCGGACCGTCTCAACACAGGTGCTCCGGTGGCGGTGGG 2475
Db 2671 AGGTTCTGCGCTCTGNAACCTTTCCAGCCGTCTTCCACAGGTCTTTCGCTGGGCGCTTGG 2730

```
QY 2476 GTCCGTCGAGATCTTCTTCAGCGCGCACAGCCCCCTGCTAGCGGTACAGAACGGCAA 2535
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2731 GTCCGTCGAAATTTCTTTCAGCAAGCACTGCCCACTTTGGTACGGATAC---GGCGCGG 2787
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2536 CCTCAAGTGGCTGAGCGCTTCGGCTATACAAACACACACATCTACCCCTTCACTCGCT 2595
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2788 GCTAAATCTCTGGAAGGTTTCTTATATCACTCCATGTTTATCCCTGGAGTTCAT 2847
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2596 CCGCTGCTCGCTTACTGCAACCTCCCGCGGCTGCTCTCACCGGCAAGTTTCAAT 2655
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2848 TCCTCTCTGCTTACTGCTACCTTGCCTGCCATCTGCTGCTCACGGGGAAGTTTATCAC 2907
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2656 GCGTCAATAGCAGTTCGCCAGCCTTCTTTCATGCCCTCTTATGATCCATCTTCGC 2715
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2908 ACAGAGCTTACCAATGTGCGCAGTATCTGGTTCATGGCACTTTTTCATCTGCTCCGT 2967
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2716 GACGGGCATCTCGAGATGCGGTGGAGCGGGGTGAGCATCAGAGAGTGTGGAGAACGA 2775
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2968 GACCGGATCTTGAATGAGTGGAGTGGCTGGCCATCGACACTGTGTGGAGAACGA 3027
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2776 GCAATTCTGGGTATCGCGCGCGGTTCGCGCATCTCTTCCCGTCTGTGAGGCGCTGCT 2835
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3028 GCAATTCTGGGTATCGCGCGCGGTTCGCGCATCTCTTCCCGTCTGTGAGGCGCTGCT 3087
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2836 CRAAGTCTCGCGGATCGACACCACTTCAACCGTCACTCCAGGCCACCGCGACGA 2895
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3088 GAAGGTGTTTCGCGGCACTGACACAGAGCTTTCACCGTGACGTCGAAGGCCGCGGACGA 3147
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2896 GGACGAGAGTTCGCCGAGCTTACGCTTCAAGTGACACAGCTCTCTATCCGCCAC 2955
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3148 G-----GAGTTCGAGCTGTACAGTTCAAGTGACACACCTCTGTATACCCCGAC 3201
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2956 CACGCTCTCATCATTAACGTCTATCGCGCTGTGGCGGATCTTCCGACGCCATCAACAA 3015
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3202 CAGCTCTCTCTGTAATTCATCGCGGTGTGGCGGATCTTCGAACGGATCAACAA 3261
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3016 CGGGTACAGTCTGGGGGCGCTTCGCGAAGCTTCTTCCGCTTCTGGGTATCGT 3075
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3262 CGGGTACAGTCTGGGGGCGCTTCGCGAAGCTTCTTCCGCTTCTGGGTATCGT 3321
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3076 CCACCTCTACCGTCTCAAGGGCTCATGGGGCGCAGAAACAGGACGCCACCGTTGT 3135
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3322 CCACCTGTACCGTCTCAAGGGCTCTGGGGGAGGAGAACAGGACGCCGATCGT 3381
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3136 TGTCAATGTTCTATCTGTGGCTCCATCTTCTCCCTGCTTGGGTGAGATCGACCC 3195
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3382 CATGCTCTGGTCCATCTCTGCGCTCGATCTTCTGCTCTGCTGCTGCGGTGCGACCC 3441
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3196 TTTTCATCTAGGACCAAGGCCCGGACGTCAGGAGTGTGGCATCAATTGC 3247
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3442 GTTCTCTCGCAAGAACAGGCCCGCTCTTGGAGAGTGTGGCTGGACTGC 3493
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 8

AAZ99509
ID AAZ99509 standard; DNA; 3813 BP.

XX AC
AAZ99509;

XX 03-JUL-2000 (first entry)

XX DNA encoding a maize cellulose synthase.

XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.

XX Zea mays.

XX Key Location/Qualifiers
XX CDS 215..3493

XX /*tag= a
XX /product= "cellulose synthase"
XX /note= "no termination codon given"

XX

PN W0200009706-A2.

PD 24-FEB-2000.

XX 16-AUG-1999; 99WO-US018760.

XX 17-AUG-1998; 98US-0096822P.

XX (PION-) PIONEER HI-BRED INT INC.

XX Dhugga KS, Helenjaris TG, Bowen BA, Wang X;

XX WPI; 2000-224343/19.

XX P-PSDB; AAY84113.

XX New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants and
PT to produce transgenic plants expressing the novel protein.

XX Claim 1; Page 129-134; 119pp; English.

XX The present sequence encodes a maize cellulose synthase polypeptide. The
CC cellulose synthase can be used for the improvement of stalk quality for
CC improved stand or silage. It also provides an increased concentration of
CC cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants and
CC seeds expressing the cellulose synthase. The polynucleotide is used for
CC modulating, preferably increasing, the level of the synthase in a plant
CC cell. The plants are preferably monocots. The polynucleotide is also used
CC as a probe or primer in the detection quantitation or isolation of gene
CC transcripts. The probes are useful in detecting deficiencies in the level
CC of mRNA in screenings for desired transgenic plant, for detecting or
CC mutations in the gene, for monitoring upregulation of expression or
CC changes in enzyme activity in screening assays of compounds, for
CC detection of any number of allelic variants of the gene, or for use as
CC molecular markers in plant breeding programs. The isolated nucleic acids
CC of their encoded polypeptides or for use as immunogens in the preparation
CC and/or screening of antibodies. The proteins can be employed in assays
CC for enzyme agonists or antagonists of enzyme function or for use of
CC immunogens or antigens to obtain antibodies specifically immunoreactive
CC with a protein

XX SQ Sequence 3813 BP; 873 A; 967 C; 1111 G; 862 T; 0 U; 0 Other;

Query Match 39.6%; Score 1364.8; DB 3; Length 3813;

Best Local Similarity 65.7%; Pred. No. 1.6e-216;

Matches 2202; Conservative 0; Mismatches 1007; Indels 143; Gaps 9;

QY 30 CTGTCGAGCACTGAGGGGTGCGAGGCGGAGAGCTAGCTAGCAGCGCGGCTCCGCGC- 88

Db 151 CTGTGCGCGCGCGGGGTTCGTGCGGAGCAGATCCGGCGGGCGGGGGGCC 210

QY 89 -GCGATGAGAGCCAGCGCGGGGTGCTGGCCGCTGCGACAAACCGAGCTGTGCT 147

Db 211 TGAGATGAGAGCTAGCGCGGGGTGCTGGCCGCTGCGATAACCGAGCTGTGCT 270

QY 148 GATCGGGGCA-CGAGGACCCCAAGCGCTGCGGGC-----CTGAGCGGGCA 195

Db 271 GATCCGCGGACCGGAGTCGGAGCCGCGGGCGGGCGCGCGCGCGGGAGGC 330

QY 196 GGTGTGCGAGATATGCGGACGAGGTTCGGGCTCAGCGTGGAGCGGACCTTCTGTCGC 255

Db 331 GCCGTGCCAGATATGCGGACGAGGTTCGGGTGGGCTTCGACGGGAGCCCTTCGTGGC 390

QY 256 CTCGAACGAGTGGGCTTCCCGTGTGCGGCCCTGTCTAGAGTACGAGCGCGGGAGGG 315

Db 391 GTGCAACGAGTGGGCTTCCCGTGTGCGGCCCTGTCTAGAGTACGAGCGCGGGAGGG 450

QY 316 CACGAGAACTGCCCCAGTGAAGACGCGCTCAAGCGCTCAAGGGAGCCCGAGGGT 375

Db 451 CTCGCAAGCGTCCCGAGTGCAGGAGCCCGCTCAAGGGCTCAAGGGTTCGCCCGGGT 510

QY	376	TGCGGGGACGATGACGAGGAGACATCGACGACTTGAGGACACGAGTTCAACATCGACGA	435	1420	CGCCATGAAGAGAGAGATGAGGAGTTCAAGGTTCGGATCAACGCCCTGGTGCCCAAGGC	1479
Db	511	GGCCGGCGACGAGGAGGAGCGCGTTCGACGACCTTGAGGGCGAGTTCCGCTCGACGA	570	1591	GGCGATGAAGAGAGATACGAGGATTTCAAGGTAAAGATCAATGCTTGGTTGCAAAAGC	1650
QY	436	CGAGATACAGCAGGCGAGCTGAGGGGACATATCGAGAACAGCCAGATCACCGAGCGAT	495	1480	CATGAGGTGCCGCGAGAGGGGTGGATCATGAGGACGSCACGCCGTGCGCCCGGGACAA	1539
Db	571	CGCGCGCGCCACGAGGACGACCGCGCATGTCGTCGCGAGTCCATGCTCAGGCGCGAGAT	630	1651	CAAAAGGTTCTTGAGGAAGGATGGCAATGCAAGATGGAAGCCCTTGCCCTTGGAAACAA	1710
QY	496	GCTGCGCGGAGGATGAGCTACGGGAGGGGCCCGCGACGAGCGGACGCGCAACACCC	555	1540	CACCCGCGACACCCCGGCATGATCCAGGTGTTCTTGGGCCACAGCGGGCGGCCACGACAC	1599
Db	631	GAGTACGGCGCGCGCGACGCGCACCCGGCTTCAGCCCGTCCCCAACGTGCGGT	690	1711	CGTACCGGATCATCTCGGAATGATTCAGGTATTCCTTGGCCAAAGTGGCGGTCTGATGT	1770
QY	556	GCAGATCCCCCCCATCATCACCGGCTCCGCTCCGTGCGGTTGAGCGGTGAGTTCCGAT	615	1600	CGAGGCGAACGAGCTGCCCGCTCGTGTACGTCTCCCTGAGAAAGGCCCGGGATTCOA	1659
Db	691	CCTCACCAACCGCCAGATGTTGATGACATCCCGCGGAGCAGCAGCGCTCGTCCGTC	750	1771	GGAAAGAAATGAGTTGCCCTCGCTTGTATGTCTCGAGAGAAAGAGGCCAGGTTATAA	1830
QY	616	TACCAAGGATAGGCGACGCGGAGGTCTGCTCTTCCCTGCAACAGGGCATCCATCGTA	675	1660	GCACCACAGAGAGCGCGGCCATGAACGCTCTGATTCGGTCTCCGCGTGTGACCAA	1719
Db	751	CTACATGAGCGCGCGCGCGCGCGGGGCAAGAGGATCCACCCGCTCCCTTTTCGAGATCC	810	1831	CCATCACAAGAAGGCTGGTGCATGAATGCATGCTGCTCTGCTGTCTTATCAAA	1890
QY	676	CCCTGTGCTGAGCCAGGAGTGCCAAAGTGGGACGAGAGAAA-----	718	1720	CGCGCCATTCATGCTCAACTTGGACTGTGATCACTACATCAACAACAGCAAGGCCATCCG	1779
Db	811	CAACCTTCAGTGAACCGAGATCCATGGACCCGTCCAGGATCTGGCCGCTACGGATA	870	1891	TGCTGCATACCTATTGAACTTGGACTGTGATCACTACATCAACATAGCAAGGCCATAA	1950
QY	719	----GAAGTGAGCTGGAAGGAGGAGATGGACGACTGGAAGTCCAAGCAGGGCATCTCT---	771	1780	GGAGGCCATGTCTTCTCATGGAACCTCAGGTCGCGCCGGAAAGGTCTGTAAGTTCAAGTT	1839
Db	871	TGGCAGCGTGGCTGGAAGGAGAGAAATGGAGGCTGGAGCGACAGACAGAGCGGCTGCA	930	1951	AGAGGCTATGTGTTTCATGATGATCCTTTGGTGGGAAAGATGTGCTATGTACAGTT	2010
QY	772	-----CGGCGCGCGCGGATCCGGAAGACATGGAAGCGCGGACGCGGACGCGGAT	819	1840	CCGCGAGAGGTTGACGCGCATCGAGTGCACGACGAGTACGCTACAGGAACACCGTCTT	1899
Db	931	GCATGTCAGAGCGAGGTTGGCGGTGATTTGGGATGGCGACGATGCGCACTAAT	990	2011	CCCTCAGAGGTTGATGATTTGACAAAATGATCGATACGCTTAAAGGAGGCTTGTCTT	2070
QY	820	CGACGAGCGAGCGAGCGGCTGTCGAGGAAGTGTGATCGCGTCGAGCAAGGTGAACCC	879	1900	CTTCGACATCAACATGAAGGGCTGGACGGCATCCAAGGCCGGGTAGCTCGGGACAGG	1959
Db	991	GGATGAGCTAGCAGGCCATTGTCCAGAAAGTCCCTATATCATCAAGCCGAATTAATCC	1050	2071	TTTTGACATCAACATGAAGGTTTGGACGGTATTCAAGGACCATTTATGTGGTACTGG	2130
QY	880	GTAACCGATGATCGTGTGCGTCTCGTTGTGCTGCGTCTTCTCCCGTACCGTAT	939	1960	GTGCGTGTTCGCGCGCAGGCGCTCTACGCGTACAACTCCCAAGGGACCCCAAGAGGC	2019
Db	1051	CTACAGGATGATATCGTTATCCGTTGGTGGTTTGGGTTTCTTCTTCCACTACCGAGT	1110	2131	ATGTGTTTTTCAGACGCGCAGGCACTGTATGTTTATGATGCTCTCTTAAACGAAGAAGCACC	2190
QY	940	CCTGCAACCCGTCGCGACGCGATCGGCTGTGGCTGTGCTCTCATCTGCGAGATCTG	999	2020	-----CAAGATGCTGACCTGCGACCTGCTGCCGCTGCTTCCGCGCAAGAGCG	2067
Db	1111	GATGCAATCCGCGGAAAGATGCAATTTGCATTTGGCTCATATCTGTAACTGTGAAATCTG	1170	2191	ATCAAGAACTTGCAACTGCTGCGCCCAAGTGGTGGCTCTCTTGTCTGCGCAGGAACAA	2250
QY	1000	GTTGCGCATCTCTGGATCTCGACCAAGTTCCCAAGTGGTTCGCCATCGACCGCGAGAC	1059	2068	GAACAACGCGCAACGCGGCTCGCGGAG-----	2095
Db	1171	GTTTGGATGCTCTGATTTCTGATCAGTTTCCCAAGTGGCTTCCATCGAGAGAGAC	1230	2251	GAATAAAGAGAGACTACAACCAAGACGGAGAGAGAAAGATATTTTTCAGAA	2310
QY	1060	GTAACCTGACCGGCTCTCTCAGGTAACGAGAGGGAAGGGAGCCGTCGCTGTGCGG	1119	2096	-----GGCACCGCTGATATGGAGT	2115
Db	1231	TTACCTGGACCGTTTGTCAAGGTTTGACAAAGGAAGGTCAACCCCTCTCAGCTTGTCTC	1290	2311	AGCAGAAAAACCCATCTCTGCGATATGCTTTTGGGTGAAATTTGATGAGGTGCTCCAGGTGC	2370
QY	1120	GGTGGACCTTTCGTGAGCAGGTGACCCGCTCAAGGAGCCCGCTGGTACCGCCAA	1179	2116	AGATAGCGACAAGGAGATGCTCATGTCCACATGAATTCGAGAGCGGTTGCGGCGATC	2175
Db	1291	AATCGACTTCTTTGTCAATAGTTCGTTGATCCCAAAAGGAACCTTCTGGTTCACGCGAA	1350	2371	TGATATCGAAGAGCGCGGAATCGTAAATCAACAGAAACTAGAGAGAAATTTTGGGCGATC	2430
QY	1180	CACGTGCTCTCCATCTCTCGCGCTGAGTACTACCCGCTGGAACAGGTCTCTGCTACGCTC	1239	2176	CGCGCGGTTTCGTCACTCGACGCTGATGAGGAAGCGCGCTCCCTCTTCTGTCGAGCCC	2235
Db	1351	CATGTCTTCCATCTTCTGTGATTTATCCGTTGAGAGGTCTCTGCTATGTTTC	1410	2431	TTCTGTTTTTGTGCGCATCAACTTCTTGAGAACGAGGGAACCTGAAGAGCGCAAGTCC	2490
QY	1240	CGACGACGCGGCTCGATGTCAGTTTCGAGTTCGCTGTGAGAGACGCGCGAGTTCGCGG	1299	2236	CGCGCGCTCTCCAAAGGAGCCATTCATGTCTAGCTCGGCTACGAGACAAGACCGA	2295
Db	1411	TGATGATGGTGTGCAATGCTTACGTTTGAAGCAATGTCTGAAACATCTGAAATTTGCAAA	1470	2491	AGCTTCTCTTCTGAAAGGAGCTATACATGTTATCAGCTCGCGCTACGAAAGACCGA	2550
QY	1300	CAAGTGGTGCCTTCTGCAAGAGTTTCGCAATGAGCCCGCGCGCGGAGTTCTACTT	1359	2296	CTGCGGCTGAGGCTGGGTTGGATCTTACGGGTGATTCACGAGGAGCATCTGACGCGGTT	2355
Db	1471	GAAATGGTTCCTTTTTCAGCAAAAAGTTTAAATATCGAGCCCTCGTCTGAGTGGTACTT	1530	2551	CTGGGGAAGAAGATTTGGCTGGATTTACCGGATCGATCAAGAGGATATCTTGACTGAT	2610
QY	1360	CTCGCTCAAGGTGCACTTACCTCAAGGACAGGTGCAAGCCACCTTCTGTCAGGAGCGCGG	1419	2356	CAAGATGCACTCCCGCGGTTGGCTCCGTGTACTGTCATGCCGAAGCGGCGGCTTCAA	2415
Db	1531	CCAACAGAGATAGACTACCTGAAGAGCAAGGTTGCTGCTTCAATTTGTTAGGAGAGGAG	1590	2611	TAAGATGCACTGCCATGGCTGGCGGTCTATTTACTGCACTCCCGAGCGGCTTGCATCAA	2670
				2416	GGGTCGCGCGCGATCAATCTATCGGACCGTCTCAACACAGGTGCTCCGCTGGCGCGTGGG	2475

QY	316	CAGCGAAGCTGCCCCAGTGCAGAGCGCTACAGCGCCTCAAGGGAGCCGAGGTT	375	1531	CCAACAGAGATAGACTACCTCAAGACAAAGTTCGTCTTCAATTTGTTAGGAGAGGAG	1590	
Db	451	CTCGAAGCTGCCCCAGTGCAGAGCGCCCGCTCAAGCGCTCAAGGGCTGCCCCGGGT	510	QY	1420	CGCCATGAAGAGAGAGTATGAGAGTTCAAGGTCCGGATCAACCGCCTGGTCCCAAGGC	1479
QY	376	TGCGGGGAGATGACGAGAGAGACATCGACGACCTTGAGACACAGATTCAACATCGACGA	435	Db	1591	GGGATGAAGAGAGATACGAGGATTCAGAGTAAAGATCAATGCCCTTGGTTCGAAAGC	1650
Db	511	GGCGGGGACGAGGAGAGACGCGCTCGAGACCTTGAGGGGCGATTGCGCTTGAGGA	570	QY	1480	CATGAAGTGCCTGGCAGAGGGGTGGATCATGAAGGACGGCACGCCGTGGCCCGGGACAA	1539
QY	436	CGAAGATCAGCAGAGGAGCTGGAGGGCAACATGCGAGAACAGCGAGATCACCGAGGCGAT	495	Db	1651	CCAAAGGTTCTCTGAGGAAGGATGGACAATGCAAGATGGAAGCCCTCGCTCGAAACAA	1710
Db	571	CGCGCGCGCCACGAGGACGACCGCAGTAGTACGTCGCGAGTCCATGCTCAGGCGCGAGAT	630	QY	1540	CACCCGCGACACCCCGCGCATGATCCAGGTGTTCTTGGGCCACAGCGCGCGCCACACAC	1599
QY	496	GCTGCGCGCAGAGATGAGTACGAGGAGGGGCCCGCAGCAGCGGCAACGCGAACAACACCC	555	Db	1711	CGTACCGCATCATCTCTGGAATGATTCAGGTATTCCTTGGCCAAAGTGGCGGTGCTGATGT	1770
Db	631	GAGTACGCGCGCGCGCGCAGCGCACCCCGGTTTCAGCCCGCTGCCCAACGTGCGCT	690	QY	1600	CGAGGCAACGAGCTGCCCGCTCGTGTACGTCTCCCGTGAGAACGCCGCCGGATTCGA	1659
QY	556	GCAGATCCCGCCCATCATCACCGGCTCCCGCTCCGTCGCGGTGAGCGGTGAGTTCCGAT	615	Db	1771	GGAGGAATGAGTTGGCTCGCTGGTATGTCTCGAGAGAAAGAGGCGCAGGTATATA	1830
Db	691	CCTCACCAACGGCCAGTGGTTGATGACATCCCGCGGAGCAGACGCGCTCGTGGCTC	750	QY	1660	GCACCAAGAAAGCGCGCCCATGAACGCTCTGATTCGCGTCTCCCGCTGCTGACCAA	1719
QY	616	TACCAACGGGTATGGCCACGCGGAGGTCTCGTCTTCCCTGCAAGCGCATCCATCCGTA	675	Db	1831	CCATCAAGAAAGGCTGGTGCCATGAATGCACTGGTCCGTCCTCTGCTCTTATCAAA	1890
Db	751	CTACATGAGCGCGCGCGCGGCGGCGCAAGAGGATCCACCCGCTCCCTTTGCGAGATCC	810	QY	1720	CGCGCATTCATGCTCAACTTGGACTGTGATCACTACATCAACACAGCAAGCCCATCCG	1779
QY	676	CCCTGTGCTGAGCCAGGAGTGCCTAGTGGGAGAGAGAA-----	718	Db	1891	TGCTGCATACCTATTGAACCTTGGACTGTGATCACTACATCAACATAGCAAGGCCATAA	1950
Db	811	CAACCTTCCAGTGCAACCGAGATCCATGGACCCGTCCTCAAGGATCTGCGCGCTACGGATA	870	QY	1780	GGAGGCATGTGCTTCTCATGGACCTCAGGTGGCGCGGAAGGTCTGCTACGTTTCAGTT	1839
QY	719	----GAAGTCAGCTGGAGGAGAGATGGACGACTGGAAGTCCAAGCAGGGCATCT---	771	Db	1951	AGAGGCTATGTGTTTCATGTGATCCCTTGGTGGGAGAAAGTGTGCTATGTACAGTT	2010
Db	871	TGCGAGGCTGGCTGGAGAGAGAAATGGAGGCTTGGAGCGTGGAGCAGAGGAGCGCTGCA	930	QY	1840	CCCGCAGAGTTTTCGACGGCATCGAGTGCACGACCGATACGCTAACAGAGAACACCGTCTT	1899
QY	772	-----CGCGCGCGCGCGCATCCCGAGACATGAGCGCGCATGTCGACTGAA	819	Db	2011	CCCTCAGAGTTTGTATGGTATTGACAAAAATGATCGATACGCTAACAGGAACTGCTCTT	2070
Db	931	GCATGTCAGGACGAGGCTGCGGTGATTGGGATGGCGACGATCGCCACTAAT	990	QY	1900	CTTCGACATCAACATGAAGGGCTGACCGCATCCAGGCCCGGTGTAAGTCCGGACAG	1959
QY	820	CGACGAGCGAGGACGCGCTGTCGAGGAAGGTGTCGATCGCTCGAGCAAGGTGAACCC	879	Db	2071	TTTTCGACATCAACATGAAGGTTTGGACGGTATTCAAGGACCCATTTATGTGGGTACTGG	2130
Db	991	GSATGAAGCTAGCGAGCCATTGTCCAGAAAGTCCCTATATCATCAAGCCGAAATTAATCC	1050	QY	1960	GTGCGTGTTCGCGCGCAGGCGCTCTACGGCTCAACCCCTCCCAAGGACCCCAAGAGGCC	2019
QY	880	GTACCGATGGTATCGTGTGCGTCTCGTTGTGCTCGCCTCTTCTCCCGTACCGTAT	939	Db	2131	ATGTGTTTTTCAGACGGCAGGCACTGTATGGTTATGATGCTCTTAAACGAGNAGCCACC	2190
Db	1051	CTACAGGATGATTATCGTTATCGGTTGGTTGGTTTGGGTTTCTTCTCAGTACCAGT	1110	QY	2020	-----CAAGATGGTGACCTGCGACTGCTGCCCGTGTCTCGCGCGCAAGAGCG	2067
QY	940	CCTGCAACCCGCTCCCGACCCATCGGGCTGTGGCTGCTCTCCATCATCTGCGAGATCTG	999	Db	2191	ATCAAGAACTTGCAACTGCTGCGCCAAAGTGGTGCCTCTCTTGTGCTGCGCAGGAGAACAA	2250
Db	1111	GATGCTCCCGCGAAGATGCAATTGCAATTGGCTCATATCTGTAATCTGTGAATCTG	1170	QY	2068	GAACACGCGCAAGGACGGCTGCCGGAG-----	2095
QY	1000	GTTGCGCATCTCTGGATCTCGACCCAGTTCCCAAGTGGTTCGCCATCGACCGCGAGAC	1059	Db	2251	GAATAAAGAGAGACTACAAACCAAGACCGAGAGAGAAAGATTTATTTTCAAGAA	2310
Db	1171	GTTCGCGATCTCTGGATTCTTGATCAGTTTCCAAAGTGGCTTCCAATCGAGAGAGAGAC	1230	QY	2096	-----GGCACCGCTGATATGGAGT	2115
QY	1060	GTACCTCGACCGCTCTCCCTCAGGTACGAGAGGAGGGAGCGCTGCTGCTGCGG	1119	Db	2311	AGCAGAAACCCATCTCTGCGCATATGCTTTGGGTGAAATTTGATGAAGGTGCTCCAGGTGC	2370
Db	1231	TTACCTGGACCGTTTGTCACTAAGTTTGAAGAGAGGTCAACCTCTCAGCTTGTCTC	1290	QY	2116	AGATAGCGACAAGAGAGATGCTCATGTCCACATGAACCTTCGAGAAAGCGGTTCGGGCGATC	2175
QY	1120	GGTGGACCTGTTCTGTGAGCAAGTGGACCGCTCAAGGAGCCGCGTGGTGAACGCCAA	1179	Db	2371	TGATATCGAAGAGGCGGGAATCGTAAATCAACAGAAACTAGAGAGAAATTTGGGCGATC	2430
Db	1291	AATCGACTCTTTGTGAGTACGTTGATCCCAAGAGGAACTCCCTTGGTTCACAGGAA	1350	QY	2176	CGCGGGTTTCGTCAAGTCCGCTGATGGAGGAAGCGCGGGTCCCTCTCTGCTCGAGCCC	2235
QY	1180	CACGCTGCTCTCATCTCCCGTAGACTACCCCGTGGACAAGGTCTCTGCTACGTC	1239	Db	2431	TTCTGTTTTTGTGCGCATCAACACTTCTTTGAGAACGAGGGAACCTTGAAGAGCGCAAGTCC	2490
Db	1351	CAGTGTCTTTTCCATCTCTGTGGATTATCCGGTTTGAAGGTCTCTCTGCTATGTTTC	1410	QY	2236	CGCGCGCTCTCAAGAGGCGCATCCATGTCTAGCTCGGCTACGAGACAAAGACCGA	2295
QY	1240	CGACGCGGCGCTGCAATGCTGAGTTCGAGTCCGCTGCGAGACGCGCGAGTTCCGCG	1299	Db	2491	AGCTTCTCTCTGAGGAAGCTTACATGTTATCAGCTCGGCTACGAGACGAGACCGA	2550
Db	1411	TGATATGGTCTGCAATGCTTACGTTTGAAGCATGTTCTGAAACATCTGAAATTTGAAA	1470	QY	2296	CTGGGGCTGAGCTGGGTGGATCTAAGGTTGATFACGAGGACATCGAGAGGATTCCTGAGCGGGTT	2355
QY	1300	CAAGTGGGTCCCTTCTGCAAGAAAGTTCCGATCGAGCCCCCGCGCGCGGAGTTCTACTT	1359	Db	2551	CTGGGAAAAGAGATTGGCTGGATTTACGAGTCGATCACAGAGGATATCTTGAAGTGGAT	2610
Db	1471	GAATGGGTTCCTTTCAGCAAAAGTTTAAATACGAGCCTCGTCTCTCGTGGTACTT	1530	QY	2356	CAAGATGCACTCCCGCGGTGGGCTCCGTTGACTGATCGAGCGAGCGGCGGCTTCAA	2415
QY	1360	CTCGCTCAAGTTCGACTACCTCAAGNACAGGTGCGAGCCACCTTCTGTCGAGAGCGCG	1419				

Db 465 GATCCGAGGGGAGGAAGGAGACGATATCTGATGCCGATGATGCTPAGCGACTTCAACTACCC 524
QY 431 -----GACGACGAGAAATCAGCAGAGGCGAGCTGGAGGGCAACATCGACAGAACGCGAGAT 483
Db 525 TGCATCTGGCAATGACGACGAGAGGAGAGATTGCTGACAGGATGCGCAGCTGGCGCAT 584
QY 484 CA-----CCGAGGGGATGCTGACGGCAGGATGAGCTACGGGAGGGGCCCCG--- 530
Db 585 GAATGCTGGGGGAGCGGGGATGTTGGCGGCCCAAGATATGACAGTGGTGAGATCGGGCT 644
QY 531 -----ACGACGGCGAGCGGCAACACACCCCGCAGATCCCGCCCATCATCACCGGCTC 582
Db 645 TACCAAGTAGCAGAGTGGTGGATCCCTCGGGGATACATCCCGTCAGTCACTAACAGCCA 704
QY 583 CCGCTCCGTCGCCGCTGAGCGGTGAGTTTCCGATTAACAACGGGTATGGCCACGCGGAGGT 642
Db 705 GATTTCGGGAGAAATCCCTGGTGCTTCCCTGACCATCATATGATGCTCTACTGGAA 764
QY 643 CTCGTCTTCCCTGCACAAGCGCATCCATCCGTACCTGTGTCTGAGCCAGGGAGTGCCAA 702
Db 765 CATTCGCGAGCGCGCCCATTTCCCTATATGAATCATTCATCAAAATCCGTCGAGGGAATT 824
QY 703 GTGGGAGGAGAAAGAGTAGCTGGAAGGAGAGATGGACGACTGGAAGTCCAGCA 762
Db 825 CTCTGTAGCGTTGGGAATGTTGCTTGGAAAGAGAGGGTTGATGGCTGGAAAATGAAGCA 884
QY 763 GGCA----- 767
Db 885 GGAACAGGGNACAATTCCTCATGACGAATGGCAACAGCATGCTCCCTGAGGGCGGGG 944
QY 768 --TCCTCGGGCGGGCCGATCCGAAAGACATGSA CGCGACGTGGCACTGAAACGCA 825
Db 945 TGTGTGTGATATTGATGTCATCAACTGATTACAACATGGAAGATGCCTTATTAAACGATGA 1004
QY 826 GCGAGGCGACCGCTGTGAGGAGAGGTGTCGATCGCGTCGAGCAAGGTGAACCCGTACCG 885
Db 1005 AACTCCCGACGCTTATCTAGGAAAGTTCCACTTCTCTCCAGGATAAATCCATACAG 1064
QY 886 GATGGTGATCGTGGTGTCTGCTGTGTGCTCGCTTCTTCTCCGCTACCGTATCTGCA 945
Db 1065 GATGGTCATTTGCTAGGATTGATTGTTCTAAGCATCTTCTTGACATACCGGATCACAA 1124
QY 946 CCCCCTCCGAGCGCCATCGGGCTGTGGTCTGCTCCATCATCTGCGAGATCTGGTTCG 1005
Db 1125 TCCTGTGCGTAATGTCATACCCTGCTGGCTTCTATCTGTTATATGTGAGATCTGCTTGC 1184
QY 1006 CATCTCTGTGATCTCGACAGTTCGCCAGTGGTTCGCCATCGACGCGGAGACGTACCT 1065
Db 1185 TCTTTCCTGGATATTGGATCAGTTCCTAAAGTGGTTTCCAAATCAACCGGAGACTTACCT 1244
QY 1066 CGACCGCTCTCCCTCAGGTACGAGAGGGAAGGGAGCGCTGCTGTGCGCGGTGA 1125
Db 1245 TGATAGACTCGCATTAAGGTATGACCGGGAGAGTGAGCCATCTCAGTTGGTCTGTTGA 1304
QY 1126 CTTGTTGTGAGACCGGTGGAACCGCTCAAGGAGCCCGCTGGTGAACGCCAACACCGT 1185
Db 1305 CATTTTGTGCTAGTGTGACCCATGAAGGAGCCTCCTTGTCTACTGCCAATACCGT 1364
QY 1186 GCTCTCATCTCGCGTAGACTACCCCGTGGAGAGAGTCTCTGCTAGCTCTCCAGCA 1245
Db 1365 GCTATCCATCTCGCTGTGGACTATCCTGTGGATAAGGTCTCTTGTATGTATCTGATGA 1424
QY 1246 CGCGGCTCGATGCTGACGTTTCAGTCTGCTGCGAGAGCGCGAGTTGCGCGCAAGTG 1305
Db 1425 TGAGCTGCTATGCTGACATTTGATGCACTAGCTGAGACTTTCAGAGTTTGTAGAAAATG 1484
QY 1306 GGTGCCCTTCTCAAGAAGTTCGGCATTCGAGCCCGCGCCCGGAGTTCTACTTCTCGCT 1365
Db 1485 GGTGCCAATTTGTAAGAAGTACAACATTAAGACCTAGAGCTCCTGAATGGTACTTCTCCCA 1544
QY 1366 CAAGGTGCACTACCTCAAGGACAAAGGTGCGAGCCACCTTGTGAGAGGGCGCGCCAT 1425
Db 1545 GAAAAATTGATTACTTGAAGGACAAAGTGCAACCTTCAATTTGTTAAAGACCGCGGGCCAT 1604

QY 1426 GAAGAGAGAGTATGAGGAGTTCAAGGTCGGATCAACGCGCTGTGGCCCAAGGCCATGAA 1485
Db 1605 GAAGAGAGATATGAAGAATTCAAATATTAGGGTAAATGGCCTTGTGCTTAAGGCAAAA 1664
QY 1486 GGTGCGGCGAGAGGGGTGATCATGAAGGACGGGACGGCGTGGCCGCGGGAACAACACCCG 1545
Db 1665 AGTCCCTGAGGAAGATGGATCATGCAAGATGGCAACCATGCCAGGAACAATACCCAG 1724
QY 1546 CGACACACCCCGCATGATCCAGGTGTTCTGTGGGCAACAGCGCGCGGCGACGACACCGAGGG 1605
Db 1725 GGACCATCTCTGAATGATTCAGGTTTTCTTGGTCAAGTGGTGTCTTGTATCTGAGGG 1784
QY 1606 CAACGAGCTGCCCGCTCGTGTAGCTCTCCGTCGAGAAGCGCCCGGGATTCAGCACCA 1665
Db 1785 TAATGAGCTACCCCGTTGGTCTATGTTCTCGTGAATAAAGCTCCTGGATTCCAGCATCA 1844
QY 1666 CAAGAAGCGCGCGCATGAACGCTCTGATTCGCTCTCCGCGCTGTGACCAACGCGGC 1725
Db 1845 CAAGAAAGCTGGTGCATGAATGCTCTTGTCCGCGTCTCAGCTGTGCTTACCAATGGACA 1904
QY 1726 ATTCATGCTCAACTTGGACTGTGATCACTACATCAACACAGCAAGGCCATCCGGGAGC 1785
Db 1905 ATACATGTTGAATCTTGATTTGATTCACATCAACACAGTAAGGCTCTCAGGAAAGC 1964
QY 1786 CATGTGCTTCTCATGGACCTTCAGGTCCGCGGAAAGGTCTGCTACGTTTCAGTTCCCGCA 1845
Db 1965 TATGTGCTTCTTATGGATCTTAACCTAGGAAGGATGTCGCTATGTTTCAGTTTCCCA 2024
QY 1846 GAGGTTGACGCGCATCGAGTCGACGACGATACGTTAAACAGGAAACACCGTCTTCTTCA 1905
Db 2025 GAGGTTGATGTTATTTGATGAGGATGATCGATATGCAACAGGAAACACCGTGTTCGA 2084
QY 1906 CATCAACATGAAGGCGCTGGACGGCATCCAAGCCCGGTGTACGTCGGGACAGGGTGCCT 1965
Db 2085 TATTAACCTTGAGAGGTCTTGATGGCATCCAAGGACAGTTTATGTGGGCACTGCTGTGT 2144
QY 1966 GTTCGGGCGCCAGGGCTCTACGGGCTCAACCCCTCCCAAGGACCAAGGCGCCCAAGAT 2025
Db 2145 TTTCAACAGAACAGCTCTATATGTTATGAGCCCCCAATTAAGCAAAAGAGGTTGTT 2204
QY 2026 GGTGACCTGCGACTGCTGCGCCCGTCTCGGCCGCAAGAGCGGAAACACGCCAAGGAC-- 2083
Db 2205 CTTGTCTCATCTATGTGTGGTGGCAGGAAGGAAAGCAATCAAGAAGGGCTTCAGACAA 2264
QY 2084 -----GGGCTGCC 2091
Db 2265 GAAAAAGTCACAGAAGCATGTGACAGTTCGTGCGCAGTATTCAATCTTGAAGATATAGA 2324
QY 2092 GGAGGGCAACCGCTGATATGGGAGTAGATAGCGACAGGAGATGCTCATGTCCCACATGAA 2151
Db 2325 GGAGGGAGTTGAAGCGCTGGATTTGATGATGAGAAATCACCTTCTTATGCTCAAAATGAG 2384
QY 2152 CTTGAGAAAGCGGTTCGGGCACTCCGCGCGGTTGCTGCTGACGCTGACGCTGATGAGGAAGG 2211
Db 2385 CTTGGAGAGAGATTGGGCAATCTGACGCTTTGTTGCTGCTCCACTCTGTAAGAAATG 2444
QY 2212 CGCGCTCCCTCTCTGTCGAGCCCGCGCGCTCCTCAAGGAGGCCATCCATGTCTATCAG 2271
Db 2445 TGGTGTCTCTAGCTTGGCACTTCCAGAAATCTCTTCTGAAGAAGCTATCAATGTCTAAG 2504
QY 2272 CTGCGGCTACGAGGACAAACCGACTGGGGGCTGGAGCTGGGGTGGATCTACGGGTCTGAT 2331
Db 2505 TTTGGCTACGAGGACAAAGATTGAATGGGGAACCTGAGATTGGGTGGATCTATGTTCTGT 2564
QY 2332 CACGAGGACATCTGACGCGGGTTCAAGATGACGCTCCCGCGGGTGGCGCTCCGTGTACTG 2391
Db 2565 GACGGAAGATAATCTCACTGGGTTCAAGATGACGCAAGGCTGGCGGTGCTGCTACTG 2624
QY 2392 CATCGGAGCGGGCGGCTTCAAGGGGTCGCGCGGATCAATCTATCGACCGCTCTCAA 2451
Db 2625 CATGCCTAAGCGCGCGCTTCAAGGGATCGGCTCCATCAATCTCTCAGACCGCTCTGAA 2684

```
QY 2452 CCAGGTGCTCGGTGGGGCTGGGGTCCGTGCGATCTTCTTACGCGGCACAGCCCCCT 2511
Db 2685 CCAGGTGCTCGGTGGGGCTTCGGTTCAAGTGAATCTTCTTACGCGGCATTTGCCCCCT 2744
QY 2512 GCTGTACGGCTTACAAAGCAACCACTCAAGTGGCTGGAGCGCTTCCGCTACATCAACAC 2571
Db 2745 ATGTTACGGGTAC---GGAGGACCCCTGAAGTCTTGGAGAGATTCCGCTACATCAACAC 2801
QY 2572 CACCATCTACCCCTTCACTCGCTCCCGCTGCTCGCCTACTGCAACCTCCCGCGCGTGTG 2631
Db 2802 CACCATCTACCCGCTCACTCGCTCCCGCTCCCTCATTTACTGTATCTGCTGCACTGTG 2861
QY 2632 CTTCTCACCAGGCAAGTTCATCATGCGCTCGATTAGCACAGTTCGCGACCTCTTCTTCAT 2691
Db 2862 CTTCTCACCAGGCAAGTTCATCATCCTCAGAGATCAGCAACTTCGCTAGTATCTGTGTTCA 2921
QY 2692 CGCCCTCTTCTATGTCCTTTCGCGAGCGGTCATCTCGAGATGCGGTGGAGCGGGGTGAG 2751
Db 2922 CTCTCTTCTCATCTCGATCTTCGCAACGGGTATCTCGAGATGAGTGGAGCGGGTGGG 2981
QY 2752 CATCGAGAGTGTGGAGGAACAGCAGTTCCTGGGTTCATCGGCGCGGTGTCGCGCATCT 2811
Db 2982 CATCGAGAGTGTGGAGGAACAGCAGTTCCTGGGTTCATCGGAGCATCTCGGCCACCT 3041
QY 2812 CTTGCGCGTGTGAGGCGCTCTCAAGTCTCTCGCGGATCGCGGATCGACCAACTTCACCT 2871
Db 3042 CTTGCGCGTCTTTCAGGCGCTCTCAAGTCTCTCGCGGATCGCGGATCGACCAACTTCACCT 3101
QY 2872 CACTCTCAAGGCGCACCGCGAGGAGGACGAGTTCGCGAGTCTTACGCTTCAAGTG 2931
Db 3102 CACTCTCAAGGCGCTG---GATGAGACGGGACTTCGCGAGTGTACATGTTCAAGTG 3158
QY 2932 GACACGCTCTTCTATCCCGCCACCAAGCTGCTCATCAATTAACGTCTACGCGCTGCGGC 2991
Db 3159 GACGACACTTCTGATCCCGCCACCAAGCTGCTCATCAATTAACGTCTACGCGCTGCGGC 3218
QY 2992 CGGCATCTCGAGCGCATCAACAGCGGTACCACTCTGCGGCGCTCTTTCGCGAGCT 3051
Db 3219 CGGCATCTCTTACGCCATCAACAGCGGTACCACTCTGCGGCGCTCTTTCGCGAGCT 3278
QY 3052 CTTCTTCGCTTCTGGGTGATCGTCCACCTCTACCGTCTCTCAAGGGGTCTATGGGGCG 3111
Db 3279 CTTCTTCGCTTCTGGGTGATCGTCCACCTCTACCGTCTCTCAAGGGGTCTATGGGGTGC 3338
QY 3112 CCAGAACAGGACGCCCGCTGTTGTTCATCTGCTGCTTCAATCTGCTGCGCTTCCATCTTCTC 3171
Db 3339 GCAGAACCGCACCCCGACCATCGTGGTGTGCTGCGGATCTCTGCTGCGCTGATCTTCTC 3398
QY 3172 CTTGCTCTGGGTAGGATCGACCTTTCATCTGTCAGGACCAAGGGCCCGGACGTGAGGCA 3231
Db 3399 CTTGCTCTGGGTTCGCTATCGATCGTTCACCAACCGGTCTACTGGCCCGGATCTCGAAC 3458
QY 3232 GTGTGGCATCAATTGCTGAGCTGT 3256
Db 3459 GTGTGGCATCAACTGCTAGGAGGT 3483
```

RESULT 11

AAZ58263

ID AAZ58263 standard; cDNA; 3776 BP.

XX

AC AAZ58263;

XX

DT 08-MAY-2000 (first entry)

XX

DE Corn cellulose synthase cDNA.

XX

KW Corn; maize; cellulose synthase; transgenic plant; ss.

XX

OS Zea mays.

XX

FH Key Location/Qualifiers

FT CDS 2..3448

FT /*tag= a
XX /partial

WO200004166-A2.

27-JAN-2000.

13-JUL-1999; 99WO-US015871.

14-JUL-1998; 98US-0092844P.

(DUPO) DU PONT DE NEMOURS & CO E I.

Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G;

PI Rafalski JA, Thorpe CJ;

XX WPI; 2000-182431/16.

DR P-PSDB; AAY58832.

XX

Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as probes for isolating cDNAs and genes encoding homologous proteins, for producing transgenic plants.

Claim 9; Page 37-38; 93pp; English.

The present sequence is that of a contig of cDNA clones and encodes a portion (see AAY58832) of corn cellulose synthase (CS). The cDNA clones were isolated from corn cob, developing tassel, seedling, shoot culture and leaf tissue cDNA libraries on the basis of homology to Arabidopsis and cotton CS sequences. The invention relates to isolated nucleic acid fragments encoding plant CS and to CS polypeptides. It also relates to the construction of a chimeric gene encoding all or a portion of the CS, in sense or antisense orientation, where expression of the gene results in altered levels of the CS in transformed host cells. The host cells can be used to screen compounds for their ability to inhibit CS activity. CS nucleic acids are also useful for producing transgenic plants having altered levels of CS, and hence altered levels of fibre. CS may also serve as a target for the development of novel herbicides

Sequence 3776 BP; 931 A; 906 C; 999 G; 940 T; 0 U; 0 Other;

Query Match

Best Local Similarity 36.8%; Score 1268; DB 3; Length 3776;

Matches 2097; Conservative 0; Mismatches 1055; Indels 126; Gaps 7;

QY 87 GCCTGATGGAGGCGAGCGCGGCTGCTGCGCGCTCGCAACACCGGAAAGAGTGTGC 146

Db 174 GGGGATGGAGGCGAGCGCGGCTGCTGCGCGCTCGCAACACCGGAAAGAGTGTGC 233

QY 147 TGATCCGGGCGCACGAGGA-----CCCAAGCGCTGCGGGCGCTGAGCGGGAGGTGT 200

Db 234 TCATCCGCGCGATGGCGAGCCAGGGCGCAAGCCCATGGACACGCGGAAAGCGCAGGTGT 293

QY 201 GCGAGATATCGCGGACGAGGTGCGGTCAAGTGGACGCGGACCTTCTGTCGCTGCA 260

Db 294 GCCAGATTTTCGCGGACGAGCTGGGGCGCAACCCCGACGGGGAGCGCTTCTGTCGCTGCA 353

QY 261 ACAGTGGCGGCTTCCCGCTGTGCGGCGCTGTCTACGAGTACGAGCGCGGGAGGCGACGC 320

Db 354 ACAGTGGCGCTTCCCGCTGTGCGGCGCTGTCTACGAGTACGAGCGCGGGAGGCGACGC 413

QY 321 AGAACTGCCCGGCTGCAAGACCGGTTCAGCGCTTCAAGGGAGCGCGGAGGTGTCCG 380

Db 414 AGAACTGCCCGGCTGCAAGACCGGTTCAGCGCTTCAAGGGAGCGCGGAGGTGTCCG 473

QY 381 GGGACCATGACGAGGAGGACATCGACGCTGCGACGCGAGTTCACATCGACGACGAGA 440

Db 474 GGGACGAGGAGGAGGACGCGCTGCGACGCTGCGAGAACCGAGTTCACATCGAGGACGAGA 533

QY 441 ATCAGCAGAGGAGGAGTTCGAGGCGCAACATGCGAGAACGAGCGAGT----- 483

Db 534 ACAGTCTCCAGTACCTTCCGCGAGTCCATGCTCCAGCCACATGAGCTACGCGCGGCG 593

Db 462 GATGACATCCGCGCGGACCGCCCTTGTGTCCTCGTTCGTTGGGTGGCGGGGGAAG 521
Qy 614 ATTACCAACGGGTATGSCCAGCGGAGGTCTCGTCTTCCCTGCAACGAGCCATCCATCCG 673
Db 522 AGGATTCAACCTCTCCGCTACGGGATCCCAACCTTCTGTGCAACGAGGTCTATGGAC 581
Qy 674 TACCTGTGTCTGAGCCAGGAGTGCAGAGTGGGACGAGAAGAAAGATGAGTCTGGAAG 733
Db 582 CCTTCAAGGATCTCGCGCATATGGCTACGGGAGCGTAGCATGGAAGGAGAGATGGAG 641
Qy 734 GAGGGATGACGACTGGAATCCNAGCAGGGGATCCTCGCGCGCGCGCGATCCGAA 793
Db 642 AGTGGGAAGCAGAGCAGGAGGATGCAACCAGACGAGGAACGATGGCGCGCGGATGAT 701
Qy 794 GACATGACCCCGACGTGGCACTGAACGACGAGCGAGGAGCGGCTGTGCGAGGAAGGTG 853
Db 702 GGTGATGATGAGATCTACCACTAATGGATGAAGTAGACAGCAATGTGCCAAGATC 761
Qy 854 TCGATCGCTGAGCAGAGGTGAACCCGATACCGATGATGATGATGATGATGATGATGATG 913
Db 762 CCGCTTCTTCAAGCCAAATCAACCCCTATAGGATGATTAATTAATTCGGCTAGTGGTT 821
Qy 914 CTGCGCTTCTTCTCGGTACCGTATCCTGCAACCCCGTCCCGACGCGATCGGGCTGTGG 973
Db 822 TTGTGTTTCTTCTTCCACTACCGAGTGAATCATCCGCTGCTGATGCAATTTGCTTATGG 881
Qy 974 CTGCTTCCCATCATCTGCGAGATCTGTTGCGCATCTCTGGAATCTCTGACCAAGTTCCCC 1033
Db 882 CTATATCTGTGATCTGTGAATTTGTTTTGCCATGTCTTGGATCTTGACCAAGTTTCCA 941
Qy 1034 AAGTGTTCCTCATCGACCGCAGAGATCACTCGACCGCTCTCCCTCAGGTACGAGAGG 1093
Db 942 AAGTGTTCCTCATCGAGAGGAAACCTATCTTGAACCGGCTGAGTTTAAGGTTTGACAAG 1001
Qy 1094 GAAGGGAGCGGTGCTGTGCGGGTGGACCTGTTCTGTGAGCAAGTGGACCGCTC 1153
Db 1002 GAAGGGGATCTTCTCAACTCGCCCTGTTGATTTCTTTGTGAGTACGGTGTGATCCCTTG 1061
Qy 1154 AAGGAGCGCGGTGGTGACCGCAACACCGTGTCTCCATCTCGCGGTGAGACTACCCC 1213
Db 1062 AAGGAACCTCANTGTGCTACTGTATATCTGTTCTATCTATCTTTCGGTGGATATCCA 1121
Qy 1214 GTGGAAGAAGTCTCTGTACTAGTCTCGACGACGCGCGTGTGATGCTGAGTCTGAGTCTG 1273
Db 1122 GTTGATAAGGTTTCATGCTACGTTTCTGATGATGTTGCTGCCATGTGTGACATTTGAAGCA 1181
Qy 1274 CTGTCGAGACGCGCGAGTTTCGCGCGCAAGTGGTGGCTTCTGTGAAGAAGTTCGGCATC 1333
Db 1182 TTGTCTGAACAACATCTGAATTTGCAAGAAATGGGTTCTTTCTGCAAAAGATATAGCCTT 1241
Qy 1334 GAGCCCGCGCCCGGAGTTCTACTTCTGCTCAAGTCTGACTACCTCAAGGACAAAGGTG 1393
Db 1242 GAGCTCTGTGCTCGAGTGTACTTCCACAGAGATAGACTACCTGAAAGCAAGGTG 1301
Qy 1394 CAGCCCACTTCTGTGAGGACGCGCGCATGAAGAGAGAGTATGAGGAGTTCAAGGTC 1453
Db 1302 GCGCCAAACTTTGTTAGAGAACCGAGAGCAATGAAGAGAGTATGAGGAAATTCAAAGGTC 1361
Qy 1454 CGSATCAACCGCTGTGTGGCCAGGCGATGAAGGTGCGCGCAGAGGGGTGGATCATGAAG 1513
Db 1362 AGAATCAATGCTTGTGTGTTAAAGCCCAAAAGGTTCTCTGAGGAAGGATGGAACAATGCAG 1421
Qy 1514 GAGCGACGCGGTGGCCGGGAAACAACCCGCGACACCCCGGCGATGATCCAGGTTTC 1573
Db 1422 GATGGAACTCCATGGCCCGGAAATATGTCCGTGATCATCTGGAATGATTCAGGTTTC 1481
Qy 1574 CTGGGCAACAGCGCGCGCAACGACACCGAGGGCAACGAGTGCCTCGCTGTGATGTC 1633
Db 1482 CTTGGTCAAAAGTGGTGGCCATGATGTGAAGGAAATGAGCTGCTCGATTTGTTTATGTT 1541
Qy 1634 TCCGCTGAGNAGCGCCCGGATTCAGCACCAAGAAAGCCGCGCGCATGAACGCTCTG 1693
Db 1634 TCCGCTGAGNAGCGCCCGGATTCAGCACCAAGAAAGCCGCGCGCATGAACGCTCTG 1693

1542 TCAAGAGAAAAACGGCCAGGCTACAACCATCAACAAGAGGCTGTGCTATGAATGCATTG 1601
1694 ATTTCGGCTCTCGCGCTGTGCTACCAACGCGCCATTCACTGCTCAACTTGGACTGTGATCAC 1753
1602 GTCCGAGTCTCTGCTGCTACTAATAATGCTCTCTATTTCGTGAACTTGGATTGTGATCAC 1661
1754 TACATCAACAACGACGAGGCCATCGGGAGGCCATGTGCTTCTCATAGCACTCAGGTC 1813
1662 TATATCAATAATAGTAAGGCTATAAGGAAGCAATGTGTTTATGATGATCCTTTGCTT 1721
1814 GCGCGGAAGGTCTGCTACGTTTCTGATCCGAGAGGTTTCGACGGCATCAGCTGACGAC 1873
1722 GGAAGAAAGTTGCTGATGTGAGTTTCTCAAGATTTGATGGGATGATGCCCATGAT 1781
1874 CGATAGCTTAAACGAGAACCCGCTCTTCTTCGACATCAACATGAAGGGCTGGAACGCATC 1933
1782 CGATATGCTTAAACGAAATGTTGCTTTTTCGATATCAACATGAAGGTTTGGATGATATC 1841
1934 CAAGCCCGGTGTACGTGCGGACAGGTCGTGTTCCGCGCCAGGCGCTCTACGCTAC 1993
1842 CAGGCGCCAAATTTATGTGGGTACTGATGTGCTTCAGAAAGCAGGCAATTATATGCTAC 1901
1994 AACCTCCCAA----- 2004
1902 GATGCTCCCAAAACAAGAAAGCCATCAAGAACTTGCAACTGTGTCGCAAAAGTGTGTC 1961
2005 -----GGGACCCCAAGAGGCCAAGATGGTGACCTGCGACTGCTGCCCG 2047
1962 ATTTGCTGTGCTGTTTGGTAAACAGGAAGACCAAGAAAGACCAAGACCTCTAAACCT 2021
2048 TGCCTTGGCCGCAAGAAG-----CGGAACACGCCAAGGACGGGTGCGGAGGCG 2098
2022 AAATTTGAGAAGATAAAGAACTTTTAAAGAAAAAGGAAAAATCAAGSCCTGTCATATGCT 2081
2099 ACGCTGATAT-----GGGAGTAGATAGACGACAGAGAGATGCTCATGTGC 2143
2082 CTTGGTGAATTTGATGAAGCCGCTCCAGGAGCTGAAAAATGAAAAAGGCTAGTATGTAAT 2141
2144 CACATGAACCTCGAGAAGCGGTTTCGGGCACTCGCGGCTTCGTCACTGCGAGCTGATG 2203
2142 CAAAGAGTTTGGAAAAAGAAATTTGGCAGTCTTCAGTTTTTGTGTCATCCACATCTT 2201
2204 GAGGAAGCGCGCTCCCTCTTGTGAGCCCCCGCGCTCTCTCAAGGAGGCAATCCAT 2263
2202 GAGAAATGGTGAACCCCTGAAGAGTGCAGTCCAGCTTCTCTTCTGAAGGAAGCTATACAT 2261
2264 GTCATCAGCTGCGGCTACGAGGACAGACCGACTGGGCTCGAGCTGGGGTGGATCTAC 2323
2262 GTCATCAGTTGTGGATATGAAGACAAAACAGGCTGGGGAAGATATTTGGTTGATTTAT 2321
2324 GGGTCGATCAAGGAGGACATCCTGACGGGTTCAAGATGCACTGCGCGGGTGGCGCTCC 2383
2322 GAATCAGTCAAGAGATATTTCTTACTGGGTTTAAAGTGCATGCCATGGTGGCGGTCA 2381
2384 GTGTACTGCATCCGAAGCGCGGCTTCAAGGGGTGCGCGCCGATCAATCTATCGGAC 2443
2382 ATTTACTGCATACCTAAACGGCGCGCTTCAAAAGTTTCGACCTCTCTCAATCTTTCCGAT 2441
2444 GCTCTCAACAGTGTCTCGGTGGGCTGGGGTTCGATGATCTTCTTCAAGCCGGCAC 2503
2442 CGTTTTCACCAAGTCTTCGCGTGGGCTCTTGGTTCAATTTGAAATTTTGTTCAGCAACCCAC 2501
2504 AGCCCTCTGCTGTACGGCTACAAGAACGGAACCTCAAGTGGCTGGAGCGCTTCGGCTAC 2563
2502 TGCCCTCTCTGGTATGGGTATGCTGGTGA---CTAAAGTTCTTGGAAAGGTTTTCGTAC 2558
2564 ATCAACACCAATCTTACCCCTTACCTCGCTCCGCTGCTCGCTTACCTGCAACCTCCCC 2623
2559 ATTAATCTCATCGTATACCTTTGGACATCTATCCGCTCTTGGCCCTTATTCGACATTCGCT 2618
2624 GCGCTGCTCTCTCACCGCAAGTTCATCATCGGCTCGATTCGATGACGCTTCGCCAGCTC 2683
2619 GCCATCTGCTGTGACAGGGGAAATTTTATACGCCAGAGCTTTAAACAATGTTGCCACCTC 2678

Db 722 ATGCATCCACTGATTACAAATATGGAAGATGCCCTTACTGATGATGAACCTCGCCAGCCTC 781
Qy 840 TGTGAGGAGAGGTGTCCATCGCTCGAGCAAGGTGAACCCGTGACCGGATGGTATCGTGG 899
Db 782 TCTCTAGAAAGTTCCCTTCCCTTCCATCCAGATAAATCCCTACAGATGGTCAATGTTTC 841
Qy 900 TGGGTCTGTTGTGTCGCCCTTCTTCTCCGTGACCGTATCCTGCAACCCGTCGCCGAGG 959
Db 842 TGGGTTTGGTGTCTAAAGCATTTTCTTCACTACCGTATTACGAATCCTGTGCGCAATG 901
Qy 960 CCATCGGGCTGTGGCTCGTCTCCATCATCTGCAGATCTGGTTGCGCATCTCTGATCC 1019
Db 902 CGTATCCGCTCTCGCTTTTACTGTATATGTGAGATTTGGTTTGTCTTGTGATAT 961
Qy 1020 TCGACCAAGTTTCCCAAGTGGTTCCCATCGACCGGAGAGTACCTCGACCGCTCTCCC 1079
Db 962 TGGATCAGTTTCCCGAAGTGGTTTCCAATCAACCGTGAACCTTACTTGTAGCTGGCAT 1021
Qy 1080 TCAGGTACGAGAGGGAAGGGAGCCGTGCTGCTGTCGTGTCGGGGTGGACCTGTTGTTGAGCA 1139
Db 1022 TGAGGTATGACAGAGAAGGTGAGCCATCTCAGTTGGCTGCTGTGACATTTTGTGAGTA 1081
Qy 1140 CGGTGACCCGCTCAAGGAGCCGCGTGTGTGACCCCAACCCGTGCTCTCCATCCTCG 1199
Db 1082 CAGTCGACCCCATGAAGGAACCTTCTTGTGTACTGCCAATCCGTGCTGTCATCTTCTG 1141
Qy 1200 CCGTAGACTACCCCGTGGACAAGGTCTCTGCTACGTCTCCGACGAGCGCGCTCGATGC 1259
Db 1142 CTGTTGATTACCCAGTGGACAAGGTCTCTTGTCTATGTATCTGACGATGGTGTGCAATGC 1201
Qy 1260 TGACGTTGAGTCGCTGTGCGAGACGGCCGAGTTTCGGGCGCAAGTGGTGCCCTTCTGCA 1319
Db 1202 TGACCTTTGATGATTTGGCTGAGACTTCAGAGTTTGTAGAAAGTGGGTTCCATTCGTTA 1261
Qy 1320 AGAAGTTTCGGACTCGAGCCCGCGCCGCGAGTTTCTCTCGCTCAAGGTCGACTACC 1379
Db 1262 AGAAGTATTAACATTGAGCCGAAGAGCTCTCAGTGGTACTTCTCCAGAAATCGATTACT 1321
Qy 1380 TCAAGGACAGGTGCGAGCCACCTTCGTGAGAGGCGCGCCATGAAGAGAGAGTATG 1439
Db 1322 TGAAGACAAAGTCCACCTTTCATTTGTTAAAGACCGTCTGTGCCATGAAGAGAGATATG 1381
Qy 1440 AGAGTTTCAAGGTCCGGATCAACGCGTGTGTCGCAAGGCCATGAAGGTGCCGGCAGAGG 1499
Db 1382 AAGAAATCAAGATTAGGATAAATGGCCTTGTCTAAGGCACAGAAAGTCCCGAGGAGG 1441
Qy 1500 GGTGGATCATGAAGGACGACCGCGTGGCCCGGGAACACACCCGACCCACCCCGGCA 1559
Db 1442 GATGGATCATGCAAGATGGCACCATGGCCAGAAACAATATAGGGACCATCTCTGGAA 1501
Qy 1560 TGATCCAGGTGTTCTTGGGCCACAGCGGCGGCCACGACACCGAGGGCAACGAGCTGCCCC 1619
Db 1502 TGATTCAAGTTTCTTGGTTCACAGTGTGCGCTTGATACCGAGGGTAATGAGCTTCCCC 1561
Qy 1620 GCCTGTGTAAGTCTCCGTTGAGAGCGCCCGGGATTCAGAGACCAAGAGAGGCGCGG 1679
Db 1562 GCCTGTGCTACGTATCTCGTGAAGCGCTCTGGGTTTCAGCACCCACAGAGAGGCTGGT 1621
Qy 1680 CCATGAACGCTCTGATTTCGCTCTCCGCTGTGCTGACCAACGCGCCATTCATCTCAACT 1739
Db 1622 CCATGAATGCTCTTGTGCTGTCTCAGCTGTGCTTACCAATGACAGTACATGTTGAATC 1681
Qy 1740 TGGACTGTGATCACTACATCAACAACAGCAAGGCCCATCCGGAGGGCCATGTGTTCTCTCA 1799
Db 1682 TTGATTGTGATCACTACATCAACAACAGCAAGGCTCTGCGGGAAGCTATGTGTTCTCTTA 1741
Qy 1800 TGAACCTCAGTGTGCGCGGAAGGTGTGCTACGTTTCAGTTTCGCGAGAGGTTTCGACGGCA 1859
Db 1742 TGGATCCAAACCTAGGAAGGAGTGTCTGTTATGTTTCAGTTTCCACAAAGGTTTCGATGGTA 1801
Qy 1860 TCGACGTGACGACCGGATCGCTAACAAGGAACCGGCTTCTTTCGACATCAACATGAAGG 1919
Db 1802 TTGATAGGAATGATCGATATGCGAACAGGAACACTGTGTTTTTCGATATTTAACTTGAGG 1861

Qy 1920 GGTGGAACGCAATCCAAAGCCCGGTGTACGTGCGGACAGGGTGCGTGTTCCGCGCGCAGG 1979
Db 1862 GTCTTGATGGCATCCAAAGGACCAAGTTTATGTGGAACTGGTGTGTGTTTCAACAGAACTG 1921
Qy 1980 CGCTCTACGGCTACAAACCTCCCAAGGAGCCCAAGA----- 2015
Db 1922 CTCATATGTTTATGAACCCCAATTTAAGCAGAAGAAGGAAGTGTTCCTTGTCTATCAC 1981
Qy 2016 -----GGCCCAAGATGGTGACTCGGACTGTGCGCGTGTCCGCGCCCAAGAACGGA 2069
Db 1982 TATGTGGGGGCAAGAAAGGCAAGCAAGTCAAAGAAAGAGTTCGGACAAGAAAGT 2041
Qy 2070 AACAAGCCAA-----GGACGGGCTGCCGAGAGGCA 2099
Db 2042 CAAACAAGCACGTGGACAGTGTGTGCCAGTGTTCATCTTGAAGATAAGAGAGGGTG 2101
Qy 2100 CGCTGATATGGAGTAGATAGGCAACAGAGATGCTCATGTCCCAATCAATCAATTCGAGA 2159
Db 2102 TTGAAGGTGCTGGATTCGATGATGAGAAATCACTTCTTATGTCTCAATGAGCTTCGAGA 2161
Qy 2160 AGCGGTTCCGGCAGTCCGCGCGTTCGTCAAGTTCGACGCTGATGAGAGGAGGCGCGTCC 2219
Db 2162 AGAGATTGCGCAGTCTGAGAGTGTGTCCTCACTCTCATGGAATATGTTGGTGTTC 2221
Qy 2220 CTCCTTGTGTCGAGCCCGCGCTCTCAAAGAGGCCATCCATGTCTCATAGCTCGGGCT 2279
Db 2222 CTCATCTCAACCCCAAGTATCTTCTTGAAGAAGCTATCCATGTGATAAGTGTGGCT 2281
Qy 2280 ACGAGCAAGACCCGACTGGGGCTGGAGC-----TGGGGTGTACT 2321
Db 2282 ATGAGGACAAGACCGAATGGGAACTGAGGCTATGCAACTTGCATGATTTGGGTGATCT 2341
Qy 2322 ACGGGTTCGATCAGGAGGACATCTCTGACGGGTTCAAAGTTCGACTCCGCGGGTGGCGCT 2381
Db 2342 ACGGTTCCGTGACAGAGATATCTCACTGGATTCAGATTCGATGCGGAGGCTGAGAT 2401
Qy 2382 CGGTGTAATGTCATGCGAAGCGGGCGGCTTCAAAGGGTTCGCGCGCATCAATCTATCGG 2441
Db 2402 CAATCTACTGATGCTCCCAAGCGCCAGCTTTCAAAGGGTCTGCTCTCATCAATCTTTTCAG 2461
Qy 2442 ACGTCTCAACCAAGTGTCTCCGTGGCGCTGGGTTCGCTCGAGATCTTCTTCAGCGCGC 2501
Db 2462 ATCGCTTAAACCAAGTGTTCGGTGGGCACCTTGGTCTCTGTGAAATCTTTTCAGTCCGC 2521
Qy 2502 ACHAGCCCTGCTGTACGCTACAGAAAGGCAACCTCAAGTGGCTGAGAGCGCTTCGCGCT 2561
Db 2522 ATGTGCCATATGGTACGGCT---ATGGAGGAGCGCTTAACTTTCGAGAGATTTGGCT 2578
Qy 2562 ACATCAACACCACTATACCTTACCTCGCTCCGCTCGCTGCTGCTACTGCAACCTCC 2621
Db 2579 ACATCAACACCACTATTTATCCATTCGATTCGATCCCGCTTCTCATATATCTGTGTTTTCG 2638
Qy 2622 CCGCGTCTGCTCTCTCAACGCGCAAGTTCATCATGCGCTCGATTCGATAGCAGCTTCGCGAGCC 2681
Db 2639 CTGCTATCTTGTGCTCACTGGGAAGTTCATCATCCAGAGATTAGCAACTTTTCTAGTA 2698
Qy 2682 TCTTCTTCACTCGCCCTCTTTCATGCTCATCTTCGAGACGGGCACTTCGAGAGATCGGTTGA 2741
Db 2699 TTTGGTTCACTCTCTCTTCAATTTTGGCACTGCTGATCTTCTTGAATGAGTGGAGTGA 2758
Qy 2742 GCGGGTGAAGCATCGAGGAGTGGTGAAGAAAGAGCAGTTCCTGGGTTCATCGGCGGGTGT 2801
Db 2759 GTGGTGTGGCATCGATGAGTGGTGGAGGAATGAACAGTTCCTGGGTTATTTGAGGATAT 2818
Qy 2802 CCGCGATCTTCTCGCGCTGTCGAGGGCTGCTCAAGTCTCTCGCGGGATCGACACCA 2861
Db 2819 CTGCGCATCTTTTGGCGCTTCTCCAGGGTCTCTCAAGGTGCTGCTGATCGACACCA 2878
Qy 2862 ACTTCAACCGTCACTCCAAAGGCGCACCGGCGAAGAGGACACAGGTTTCGCGAGCTCTACG 2921
Db 2879 ATTTCACTGTCACTCAAGGCTTCTGATGAGATGCGGAC---TTTGTGAGCTCTTACA 2935

QY 2922 CCTCAAGTGGACCAAGCTCTCTATCCCGCCACCAAGCTCTCTATCATTAACGTATCG 2981
 Db 2936 TGTTCAGTGGACCAAGCTCTCTATCCCGCCACCAAGCTCTCTATCATTAACGTATCG 2995
 QY 2982 GCGTCGTGGCGGCGATCTCCGACGCGCATCAACAAACGGGTACCACTCTGGGGGCGCCCTCT 3041
 Db 2996 GTGTTGTCTGGTATCTCATACGGGATCAACAGGGGTACCACTCTGGGGGCGCCCTCT 3055
 QY 3042 TCGCAAGCTCTCTCTGGCTTCTGGGTCATCGTCCACCTCTACCGGTTCTCAAGGGGC 3101
 Db 3056 TTGCAAGCTCTCTCTGGCTTCTGGGTCATCGTCCACCTCTACCGGTTCTCTGGCTT 3175
 QY 3102 TCATGGGCGCCAGACAGGAGCCGACCGTGTGTCTCATCTGTCCATTTCTGGCTT 3161
 Db 3116 TTATGGGTCGCAAAACCGCACTCCGACCATCTGTGTGTGTGGGCAATCTCTGGCTT 3175
 QY 3162 CCATCTTCTCTCTCTCTGGGTCAGGATCGACCTCTTCATCTGTCAGGACCAAGGCGCGG 3221
 Db 3176 CGATCTTCTCTCTCTCTGGGTCAGGATCGACCTCTTCATCTGTCAGGACCAAGGCGCGG 3221
 QY 3222 ACGTCAGGAGTGTGGCATCAATTTGCTGA 3250
 Db 3236 ATACCCAAACATGTGGCATCACTGCTAA 3264

RESULT 14

AAZ99533
 ID AAZ99533 standard; DNA; 3704 BP.

XX AAZ99533;

DT 03-JUL-2000 (first entry)

XX DNA encoding a maize cellulose synthase.

XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
 KW transgenic plant; plant breeding marker; ss.

XX Zea mays.

XX Key Location/Qualifiers

FT 272..3496

FT /*tag= a

FT /product= "cellulose synthase"

FT /transl_except= (pos: 1544..1546, aa: Xaa)

FT /note= "no termination codon given; Xaa is an unspecified
 amino acid"

XX W020009706-A2.

XX 24-FEB-2000.

XX 16-AUG-1999; 99WO-US018760.

XX 17-AUG-1998; 98US-0096822P.

XX (PION-) PIONEER HI-BRED INT INC.

XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;

DR WPI; 2000-224343/19.

DR P-PSDB; AAY84121.

XX New genes which encode maize cellulose synthase polypeptides in plants
 PT useful for modulating the expression of cellulose synthase in plants and
 PT to produce transgenic plants expressing the novel protein.

XX Claim 1; Page 191-196; 119pp; English.

XX The present sequence encodes a maize cellulose synthase polypeptide. The
 CC cellulose synthase can be used for the improvement of stalk quality for
 CC improved stand or silage. It also provides an increased concentration of
 CC cellulose in the pericarp, hardening the kernel and improving its

CC handling ability. The sequences are used to produce transgenic plants and
 CC seeds expressing the cellulose synthase. The polynucleotide is used for
 CC modulating, preferably increasing, the level of the synthase in a plant
 CC cell. The plants are preferably monocots. The polynucleotide is also used
 CC as a probe or primer in the detection quantitation or isolation of gene
 CC transcripts. The probes are useful in detecting deficiencies in the level
 CC of mRNA in screenings for desired transgenic plant, for detecting or
 CC mutations in the gene, for monitoring upregulation of expression or
 CC changes in enzyme activity in screening assays of compounds, for
 CC detection of any number of allelic variants of the gene, or for use as
 CC molecular markers in plant breeding programs. The isolated nucleic acids
 CC of the present invention can also be used for recombinant expression of
 CC their encoded polypeptides or for use as immunogens in the preparation
 CC and/or screening of antibodies. The proteins can be employed in assays
 CC for enzyme agonists or antagonists of enzyme function or for use of
 CC immunogens or antigens to obtain antibodies specifically immunoreactive
 CC with a protein

XX
 SQ Sequence 3704 BP; 909 A; 838 C; 962 G; 994 T; 0 U; 1 Other;

Query Match 35.1%; Score 1207.4; DB 3; Length 3704;
 Best Local Similarity 63.6%; Pred. No. 1.8e-190;
 Matches 2076; Conservative 0; Mismatches 1047; Indels 141; Gaps 10;

QY 155 GGCACGAGACCCCAAGCCGCTGCGGGCGCTGAGCGGGCAGGTGTCGAGATATGCGGC 214
 Db 278 GCGCGCGACGCCACGATTCGGGGAAGCATGTGCGCGGCGAGGTGCCAGATTCGCGC 337
 QY 215 GACGAGTGGGCTACGGTGGACGGGACCTCTTCGTCGCTCAACGAGTGGGCTTC 274
 Db 338 GACGGGTGGGACCGCGCGGACCGGACCTCTTCACCGCTCGACGCTTCGCGCTTC 397
 QY 275 CCGGTGTCGGCGCCCTGCTACGAGTACGAGCGCGGGAGGGCAGCGAGAACTGCCCCAG 334
 Db 398 CCGGTGTCGGCGCCCTGCTACGAGTACGAGCGCAAGGACGCCACCCAGGCGTGC 457
 QY 335 TGCAGCGCGCTACAGCGCCTCAAGGGAGCCCGAGGTTGCCGGGACGATGACGAG 394
 Db 458 TGCAGAACTAAGTACAGCGCCCAAGAGGGAGGCCACCACTACACGCTGAGGAAATGAG 517
 QY 395 GAGGACATCAGCAGCTGGAGCAGGTTCAACATC-----GACGACGAGAAATCAG 445
 Db 518 GATGTGATGCTGACGATGTGAGTACTACACTACCAAGCATCTGCAACCAAGATCAG 577
 QY 446 CAGAGGCGCTGGAGGCAACATGCAAG-----ACAGCCAGATCACCGAGCGATGCTGCA 501
 Db 578 AAGCAAAAGATTGCTGAGAGAACTGCTCACTTGGCGGACAACTCACGCTGGCAGTGATTT 637
 QY 502 CGGAGGAT-----GAGCTACGGGAGGGGCCCCGACGAGCGGCGACGGCAACAACA 551
 Db 638 GGCCTGGCTAAGTATGACAGCGGTGAAATTTGGGCATGGGAAGTATGACAGTGTGATC 697
 QY 552 CCGCGCAGATCCCGCCCATCATCAGCGCT--CCCGCTCCGTCGCGGTGAGCGTGT 609
 Db 698 CCTGTGGATATATCCCGTCACTAACTATAGCAGATCTCAGGAGAGATTCCTGGAGCT 757
 QY 610 TCCGATTACCAACGGGTATGCGCCAGCGAGGTCTCGTCTTCCCTGCACAAAGCGCATCCA 669
 Db 758 TCCCTGATCATATGATGCTCTCTGTTGGGAACATTTGGCAGGCTGGACATCAATTTCT 817
 QY 670 TCCGT--ACCTGTGTCTGAGCCAGGAGTGCCAAAGTGGGACGAGAAAGAGTGAAGC 727
 Db 818 TATGTAAATCATTTCTCCAAACCCATCGAGGAGGTTCCTCCGCTAGCCTTGGCAATGTTGCA 877
 QY 728 TGGAGGAGGAGGATGACGACTCGAAGTCCAA----- 759
 Db 878 TGGAAAGAGAGGGTGGATGGATGGAATAAGGATAAAGGTGCAATTCCTATGACCAAT 937
 QY 760 -----GCAGGCGATCTCTCGCGCGCGCGCGCGATCCCGGAA 793
 Db 938 GGAACAAGCATTGCTCCATCAGAGGCGGTGGAGTGTCTGATATTGATGCTTCTACTGAT 997
 QY 794 GACATGACCCCGACGCTGGCACTGAACGACGAGGCGGAGGCGCGCTGTGCGAGGAAGTG 853

Db 998 TATAAATGAGATGCTTACTGAAATGATGAACCTCGGCAACCTCTATCTAGAAAGTG 1057
QY 854 TCATCGCGTCGAGCAAGGTGAACCCGTACCGGATGGTGTGATCGGTGGTGTGCTCGTGTGG 913
Db 1058 CCATTTCTTTCATCCAGATAAATCCGTACAGAAATGGTCATGTGTCTAGCTTGGCTGTT 1117
QY 914 CTCGCTTCTTCTCCGCTACCGTATCTCTGTCACCCCGTCCCGACGCCATCGGGCTGTGG 973
Db 1118 CTATGCATATTTCTTGGCTACCGTATCACACATCTCTGTGAACAAATGCATATCCACTGTGG 1177
QY 974 CTCGCTCCATCATCTCGAGATCTGGTTCGCCATCTCTCTGGATCTCTGCACCAAGTTCCCC 1033
Db 1178 CTTTATCCGTATATGTGAGATCTGGTTGGCTTTGTCTGGATTTTGGATCAGTTCCCA 1237
QY 1034 AAGTGTGTTCCCATCGACCGCAGAGGTACTCTCGACCGCTCTCCCTCAGGTACGAGAGG 1093
Db 1238 AAGTGTGTTCCCATCAACCGTGAAACATACCTTTGATAGACTGGCTTTAAGGTATGACCGA 1297
QY 1094 GAAGGGAGCCGTGCTGTCTGTCGCGGGTGGACCTGTTCTGTGAGCAGCGTGGACCCGCTC 1153
Db 1298 GAAGGTGAACCATCTCAATTAGCTCTCTGTGTGATATTTTGTCACTACTGTGGATCCAATG 1357
QY 1154 AAGGACCGCGCTGTGTGACCGCAACACCGTCTCTCCATCTCTCGCGTAGACTACCCC 1213
Db 1358 AAGGACCTCCTTGTCACTGCAAAATACTGTGCTTTTCCATCTCTGTCGATTAATCCG 1417
QY 1214 GTGCAAGGTCTCTGCTACGTCTCCGACGCGCGCTCGATGCTGACGTTTCGAGTGG 1273
Db 1418 GTTGACAAAGTATCTGCTATGTTTCGGATGATGGAGCTGCTATGCTGACTTTTGATGCT 1477
QY 1274 CTGTGCGAGACGCGCGAGTTCCGCGCAAGTGGTGCCTTTCTGCAAGAAAGTTTCGCAATC 1333
Db 1478 CTCTCTGAAACTCAGAGTTGTAGAAAATGGTTCCGTTCTGTGAAGATACAACATA 1537
QY 1334 GAGCCCGCGCCCGGAGTTCTACTTTCTCGCTCAAGTTCGACTACCTCAAGGACAGGTG 1393
Db 1538 GAGCCTTANGCCCGCGAAATGGTACTTTGTCTCAAAAATTTGATTTACTTGAAGACAAAGTT 1597
QY 1394 CAGCCCACTTCTGTGAGAGCGCGCGCATGAAGAGAGATGATGAGGAGTTTCAAGTTC 1453
Db 1598 CAACCTCATTTGTGAAGAACGCGCGGCCATGAAGAGAGATATGAAGNAATCAAGTT 1657
QY 1454 CGGATCAACGCGTGTGTGCAAGGCCATGAAGGTCCCGCAGAGGGGTGGATCATGAAG 1513
Db 1658 CGTATCAATGGTCTTGTAGCCAGGCACAAAAGTTTCCGAGGAGGATGGATCATGCAA 1717
QY 1514 GACGCAACGCGTGGCGCGGACACACCCGCGACACCCCGCATGATCCAGGTGTTTC 1573
Db 1718 GATGGTACACCTTGGCTTGGGAACAATACTAGGGACCATCTCGGAATGATTCAGGTTTTC 1777
QY 1574 CTGGGCGCACAGCGCGGCACAGCACCGAGGGCAACGAGCTGCGCCGCTCGTGTACGTC 1633
Db 1778 CTGGGTCAAGTGGAGGGCTTGACGTTGAAGGCATGAACCTTCTCGTTGGTTATGTG 1837
QY 1634 TCCCGTGAGAAGCGCCGGATTCAGCACCAAGAAAGCCGCGCCATGAACGCTCTG 1693
Db 1838 TCTCGTGAATAAAGCTCTCTGATTTCCAACATCAAGAAAGGCTGGTCCCATGAATGACATT 1897
QY 1694 ATTCGCTCTCCCGCTGCTGACCAACGCGCCATTATGCTCAACTTGAAGTGTGATCAC 1753
Db 1898 GTTCGTGTATCAGCTGTCTTACTTAATGGGCAATACATGTTGAATCTTGAATGTGACCAC 1957
QY 1754 TACATCAACAAAGCAAGGCCATCTCGGGAGCCATGTGTTCTCTCATGACCCCTCAGGTTC 1813
Db 1958 TACATCAATATAGCAAGGCTCTTCGAGAAGCTATGTGTTCTTATGAGACCCAAACCTA 2017
QY 1814 GGCCGAAGGTCTGTCTACGTTTCAGTTCCGCGAGAGGTTTCGACGCGATCGACGTGACGAC 1873
Db 2018 GGAAGGAATGTCTGTTATGTGCCAAATTTCTTCAGAGGTTTGAATGTTATGATGAATGAC 2077
QY 1874 CGATCGCTTAACAGGAACACCGTCTTCTTCGACATCAACATGAAGGGCTGGACCGCATC 1933

Db 2078 CGATATGCAACAGGAACACTGTGTTTTTCGATATAATTAACCTGAGAGGTCTTGACGCCATT 2137
QY 1934 CAAGGCCCGGTGTAGCTCGGACAGGGTGGTGTTCGCGCGCAGCGCTCTACGCGCTAC 1993
Db 2138 CAAGGCCCAAGTTTATGTGGAACTGGTTGTGTGTTTAAACAGAACGCGCTTATATGTTAT 2197
QY 1994 AACCTTCCAAAGGACCCCAAGAGGCC-----CAAGATG 2026
Db 2198 GAGCCTCCAGTCAAGAAAAAAGCCAGGCTTCTTCTCTTCGCTTGTGGGGGAAGAAA 2257
QY 2027 GTGACCTCGCATCTGCTGCCGCTTCGCGCCGACAGAGGGGAACACGCCAAGGAC--- 2083
Db 2258 AAGACGTCAAAATCTAAGAAGAGCTCGGAAAAAGAAAGATACATAGACAGCAGACAGT 2317
QY 2084 -----GGGCTGCCGAGGGCACCGCTGATATGGAGTAGAT 2119
Db 2318 TCTGTACCAAGTATTTAACTCTCGAAGATATAGAGGAAGGATTTGAAGTTTCTCAGTTTGT 2377
QY 2120 AGCGACAAAGGATGCTCATGTCCCAATGAATCTTCGAGAAGCGGTTTCGGGCAAGTCCGGG 2179
Db 2378 GATGAGAAATCGCTGATTATGTCTCAAAATGAGCTTGGAGAAGAGATTTGGCCAGTCCAGT 2437
QY 2180 GCGTTGCTCACGTCGACGCTGATGAGGAGGGCGGTCCCTCTTCGTCGAGCCCGCC 2239
Db 2438 GTTTTGTAGCCTCTACTCTGATGGAATATGGTGGTGTCCACAATCTGCAACTCCAGAG 2497
QY 2240 GCGCTCCTCAAGAGGCCATCCATGTCTCAGCTCGGCTAGCAGGACAAAGACCGACTGG 2299
Db 2498 TCTCTTCTGAAGAAGCTATTTCAATGTCTCAGCTGTGGCTATGAGGACAAAAGTGACTGG 2557
QY 2300 GGGCTGGAGCTGGGTGGATCTACGGGTGATACGCGTCGATACGAGAGCAATCTCTGACGGGTTCAAG 2359
Db 2558 GGAACCTGAGATTTGGGTGGATCTATGGTCTGTGTACAGAAAGACATCTCACCAGGATTTCAAG 2617
QY 2360 ATGCACTGCGCGGGTGGGCTCCGTGTACTGCATCGCCGAGCGGGCGGGTTCAGAGGG 2419
Db 2618 ATGCACTGTCGAGGCTGGCGATCAATCTACTGCATGCCCTAAAGCAGCAGCTTTCAAGGGA 2677
QY 2420 TCGGGCCGATCAATCTATCGGACCGTCTCAACACGAGTCTCCGGTGGCGCTGGGGTCC 2479
Db 2678 TCTGCTCCTATCAACCTTTCCGATCGTTTGAATCAAGTCTTCGGTGGGCTCTTGGTTCC 2737
QY 2480 GTCGAGATCTTCTTACGCGGCACAGCCCCCTGTGTGTAAGGCTACAAAGAACGGCAACCTC 2539
Db 2738 ATTTGAAATTTCTTTTTCAGCAGGCAATGTCCCATATGGTATGGCT---ATGAGAGCCGCGCTT 2794
QY 2540 AAGTGGCTGGAGGCTTGGCTTACATCAACACCATCTACCCCTTCACTCCGCTCCCG 2599
Db 2795 AAATTCCTGGAGAGATTTGCTTTATATCAACAACAATTTATCCACTCAATCAATCCCG 2854
QY 2600 CTGCTCGCTACTGCAACCTCCCGCGTCTGCTCCCTCACCGGCAAGTTTCATCATGCGG 2659
Db 2855 CTCTCTCTGCTACTGCATATTCAGCAGAGTTTGTCTCTCAGTGGGAAGTTTCATCATCCCA 2914
QY 2660 TCGATTAGACGTTTCGCCAGCCTCTTCTTCATCGCCCTCTTCATGTTCATCTTCGCGAGC 2719
Db 2915 AAGATTAGTAACTAGAGAGTGTGTGGTTTATATGCTCTTTATCTCAATCTTTGCACT 2974
QY 2720 GGCATCTCGAGATCGGTGGAGCGGGTGAACATCGAGGATGGTGGAGGACGAGCAG 2779
Db 2975 GGTATCTTTGATGAGGTGGAGTGGTGTGGCATTTGATGAATGGTGGAGGACGAGCAG 3034
QY 2780 TTTCTGGGTCTATCGGGCGGTGTCCGCGATCTCTTCGCGCTCGTGCAGGGCTCTGCTCAAG 2839
Db 3035 TTTCTGGGTCTATTTGGTGTATTTCTGGGCATTTATTTGGCGCTTTCAGAGGTCTCTCTGAAG 3094
QY 2840 GTCTCTCGCGGATCGAACCAATTTACCGTCACTCTCAAGGCCACCGCGGACGAGGAC 2899
Db 3095 GTGCTTGTGTGTATCGACACGAGCTTCACTGTCACTCTTAAGGCCAC---TGACGAAGAA 3151
QY 2900 GACGAGTTTCGCGAGCTCTACGCTTCAAGTGGACACCGCTCTCTATCCCGCCACCG 2959
Db 3152 GGTGATTTTGGCCAGCTCTACATGTTCAAGTGAGAACCGCTTCTGTATCCACCACT 3211

Db	3192	TTTGGGAAGCTATTCTTTTGCAATTTTGGTGATTGTGTCATCTTTATCCCTTTCTCAAAGGT	3251
Qy	3101	CTCATGGGGCGCCAGAAACAGGACGCCACCGTTGTTGTTCATCTGCTCCATTCTGCTGGCC	3160
Db	3252	TTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCT	3311
Qy	3161	TCCATCTTCTCCCTGCTCTGGGTGAGGATCGACCCCTTTTCATCGTCAGGACCAAGGGCCCG	3220
Db	3312	TCAATCTTCTCGCTCCTTTGGGTTTCGGATTGATCCTTTCCCTTGGGAAGGATGATGGTCCG	3371
Qy	3221	GACGTACGGCAGTGTGGCATCAATTGC	3247
Db	3372	CTTCTTGAGGAGTGTGTTGGATTGC	3398

Search completed: March 11, 2005, 19:16:34
Job time : 1207 secs

This Page Blank (uspto)

[illegible]

```
QY 2457 TGCTCCGGTGGCGCTGGGCTCCGTCGAGATCTTCTTCAGCCGCGCAGACCCCTCTGCTGT 2516
Db |||||||
QY 2467 TGCTCCGGTGGGCTCTTGGGTCCTGGAGATCTTCTTCAGCCGCGCATGCCCCCTGTGGT 2826
Db |||||||
QY 2517 ACAGCTACAAGAACGCGCAACTCAAGTGGCTGGAGCGCTTCGCTACATCAACACCAACCA 2576
Db |||||||
QY 2827 ACAGCTAC---GGAGGGCGGCTCAAGTTCCTGGAGAGATTCGCTACATCAACACCAACCA 2883
Db |||||||
QY 2577 TCTACCCCTTCACTCCGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2636
Db |||||||
QY 2884 TCTACCCGCTCACCTCCATCCGCTTCTCTCATCTACTGCTGCTGCTGCTGCTGCTGCTGCT 2943
Db |||||||
QY 2637 TCACCGGCAAGTTCATCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2696
Db |||||||
QY 2844 TCACCGGAAAGTTCATCATTCAGAGATCAGCAACTTCGCGAGCATCTGGTTCATCTCCC 3003
Db |||||||
QY 2697 TCTTCATGCTCATCTTCGCGACGGGCTCCTGGAGATGCGGTGGAGCGGGTGAAGCATCG 2756
Db |||||||
QY 3004 TCTTCATCTCGATCTTCGCCACGGGATCTGGAGATGAGTGGAGCGGGTGGGATCG 3063
Db |||||||
QY 2757 AGAGTGGTGGAGAAACAGCAGTTCCTGGGTCATCGCGCGGCTGCTCGCGCATCTCTTCG 2816
Db |||||||
QY 3064 ACAGTGGTGGAGAAACAGCAGTTCCTGGGTCATCGGGGCTCTCGCGCATCTCTTCG 3123
Db |||||||
QY 2817 CGGTGCTGACGGGCTGCTCAAGTCTCTCGCGGGATCGACCAACTTCACCGTCACTT 2876
Db |||||||
QY 3124 CCGTGTTCAGGGCTGCTCAAGTGTCTGGCGGCTCGACCAACTTCACCGTCACTT 3183
Db |||||||
QY 2877 CCAAGGCCACCGGACGAGACGAGTTCGCGGAGCTCTACGCTTCAAGTGGACCA 2936
Db |||||||
QY 3184 CCAAGGCTCG---GACGAGACGGGCTCTCGCGGAGCTGTATGTTCAAGTGGACGA 3240
Db |||||||
QY 2937 CGTCTCTATCCCGCCACCAACGCTGCTCATCTTAACGTCATCGGGTCTGTCGCGGCA 2996
Db |||||||
QY 3241 CGTCTCTATCCCGCCACCAACGCTGCTCATCTTAACGTCATCGGGTCTGTCGCGGCA 3300
Db |||||||
QY 2997 TCTCGAGCGCATCAACAAACGGGTACAGTCTCGGGGCGCTCTTCGGCAAGCTCTTCT 3056
Db |||||||
QY 3301 TCTCTACGCGCATCAACAGCGGATACAGTCTGGGGCGGCTCTTCGGCAAGCTCTTCT 3360
Db |||||||
QY 3057 TCGCTTCTGGTTCATGCTCCACCTCTACCGCTTCTCAAGGGGCTCATGGGGCGCAGA 3116
Db |||||||
QY 3361 TCGCTTCTGGTTCATGCTCCACCTGTACCGCTTCTCAAGGGGCTCATGGGGCGCAGA 3420
Db |||||||
QY 3117 ACAGAGCCCGCACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3176
Db |||||||
QY 3421 ACAGAGCCCGCACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
Db |||||||
QY 3177 TCTGGGTGAGATCGACCTTTCATGCTCAGGACCAAGGGGCGGAGCTCAGGCGAGTGTG 3236
Db |||||||
QY 3481 TGTGGGTTCGATCGACCGCTTCCACCACCGGCTCACTGGCGCGGATACCCAGAGCTGTG 3540
Db |||||||
QY 3237 GCATCAATGCT 3248
Db |||||||
QY 3541 GCATCAATGCT 3552
Db |||||||
```

RESULT 2

```
US-10-160-719A-49
; Sequence 49, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
```

```
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 3746
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (321)...(3449)
US-10-160-719A-49
```

Query Match

```
Best Local Similarity 40.6%; Score 1397.6; DB 4; Length 3746;
Matches 2185; Conservative 0; Mismatches 924; Indels 143; Gaps 9;

QY 133 GAACGAGCTGCTGATCCGGGCGCACGAGGACCCCAAGCGCTCGGGCGCTGAGCGG 192
Db |||||||
QY 193 GCAGGTGTGCGAGATATGCGCGCACGAGGTGCGGCTCACGCTGCGACCGCGCTCTTCGT 252
Db |||||||
QY 368 ACAGGTGTGCGAGTCTGCGCGCACGCGGTGGGACCAACCGCGGAGGGGACGCTCTTCG 427
Db |||||||
QY 253 CGCTGCAACGAGTGTGCGGCTTCCCGGTGCTCGGCTCTGCTACGAGTACGAGCCCGGGA 312
Db |||||||
QY 428 CGCTGCGAGCTCTGCGGTTTCCGGTGTGCGGCTCTGCTACGAGTACGAGCGCAAGGA 487
Db |||||||
QY 313 GGCACGCGAATCTGCCCCAGTGCAGACGCGCTACAGGCTCAAGGGGAGCCGAG 372
Db |||||||
QY 488 CGGACGCGAGGCTGCGGCGGCTGCAAGCAAGTACAGCGCCACAGGGGAGCCCGG 547
Db |||||||
QY 373 GGTTCGCGGCGAGTACGAGGAGGACATCGACGCTCGAGCACGAGTTCACACATCGA 432
Db |||||||
QY 548 GATCCGTGGGAGGAGGAGACGACACTGATGCGATAGCGACTTCAATTACCTTGATC 607
Db |||||||
QY 433 CGACGAGAAATCAGCAGAGGAGTGGAGGCAACATGCAAGACAGCAGATCA----- 485
Db |||||||
QY 608 TGGCAATGAGGACCAAGAGCAGAAATGCGCAGAAATGCGAGCTGGGCGCATGAACGT 667
Db |||||||
QY 486 -----CCGAGGCGATGCTGCGAGGAGGATGAGTACGCGGAGGGGCGCCG----- 530
Db |||||||
QY 668 TGGGGGCGAGCGGGAGTGTGTCGCCCAAGTATGACAGTGGCGAGATCGGGCTTACCAA 727
Db |||||||
QY 531 --ACGAGCGAGCGGCAACCAACCGCGAGATCCCGCCCATCATCAGCGCTCCCGGCTC 588
Db |||||||
QY 728 GTATGACAGTGGCGAGATTCCTCGGGATATATCCCATCAGTCACTAACGCCAGATCTC 787
Db |||||||
QY 589 CGTCCCGGTGAGCGGTGAGTTTC--GATTACCAACGGGTATGGCCACGCGGAGGTCTCGT 647
Db |||||||
QY 788 AGGAGAAATCCCTGGTCTTCCCTGACCATCATATGATGTCCTCCCACTGCGAAACATGG 847
Db |||||||
QY 648 CTTCCCTGCAACGCGCATCCATCCGTGCTGCTGAGCCAGGAGTGCAGAGTGGG 707
Db |||||||
QY 848 CAAGCGTCTCCATTTTC--CCTATGTGAACCATTCGCCCAATTCGCTGAAGGAGTCTCTG 906
Db |||||||
QY 708 ACAGAGAAAGAAAGTGAAGTGAAGAGGATGAGCAGCTGGAAGTCCAAAGCAGGCA 767
Db |||||||
QY 907 GTAGCATTTGGGATTTGCTTGGAAAGAGAGGTTGATGCTGGAAATGAGAGCAGACA 966
Db |||||||
QY 768 -----TCC 770
Db |||||||
QY 967 AGGGAGCATTTCCCATGACGAATGGCAAGCATTTGCTCCTCTGAGGGTGGGGTGTG 1026
Db |||||||
QY 771 TCGCGCGCGCGCGATCCCGAAGACATGAGCGCGGAGCTGGGCTGAAACGACGAGGCA 830
Db |||||||
QY 1027 GTGATATTGATGATCAACTGATTACCAACATGGAAGATGCTTTATTGAACGACGAAATC 1086
Db |||||||
QY 831 GGCAGCGCTGTCGAGGAAGTGTGATCCGCTGAGCAAGGTGAACCCGCTACCGGATGG 890
Db |||||||
QY 1087 GACAGCCTCTATCTAGGAAAGTTCACCTTCTCTCCAGGATTAATCCATACAGATGG 1146
Db |||||||
```


Db 1704 GAGAAATATGAGAAATTCANAAGTTAGGGTAAATGSCCTTGTGTGCTAAGGGACACAGAAAGTTC 1763
Qy 1491 CGGCAGAGGGGTGGATCATGAAGGACGCGCGTGGCCCGGGAAACAAACACCGCGAAC 1550
Db 1764 CTGAGGAGAGATGGATCATGAAGATGGACACCATGGCCAGGAAACAAATACAGGGACC 1823
Qy 1551 ACCCGGCATGATCCAGGTGTTCTGGGCAACAGCGCGGCCACGACACCGAGGGCAAG 1610
Db 1824 ATCTGGAATGATTCAGGTTTCTGTTGTCACAGTGGTGGCTTGATCTACTGAGGGCAATG 1883
Qy 1611 AGTGGCCCGCTCGTGTAGTCTCCCGTGAGAGCGCCGGGATTCACACACACAGA 1670
Db 1884 AGCTACCCCGTTTGGTCTATGTTTCTGTAAGAGCGTCTGGATTCCAGCATCACAGA 1943
Qy 1671 AGCGCGCGCATGAACGCTGTGATTCGCTGCTCCCGCGTGTGACCAACGCGCATTC 1730
Db 1944 AGCTGGTGCAATGAATGCTCTGTTGCTGCTCAGCTGTGCTTACCAATGGACATACA 2003
Qy 1731 TGCTCAACTTGGACTGTGATCACTACATCAACACAGCAAGGCGCATCCGGAGGCCATGT 1790
Db 2004 TGTGAACTTGTGATGATCACTACATTAACACAGTAAGGCTCTCAGGGAAGCTATGT 2063
Qy 1791 GCTTCCTCATGGAACCTCAGTCCGCGGAGGTCTGCTAGTTCAGTTCGCGCAGAGGT 1850
Db 2064 GCTTCCTTATGGACCCCTAAGCTTAGGAAGGAGTGTCTGCTACGTCAGTTCCTCCAGAGAT 2123
Qy 1851 TCGACGCGCATCGAGCTGACGCGATACGCTTAACAGGAACACCGTCTCTTCTCGACATCA 1910
Db 2124 TCGATGGCATGACAGGAATGATCGATATGCCAACAGGAACACCGTGTGTTTTCGATATTA 2183
Qy 1911 ACATGAAGGGCTGGACGGCATCAAGCGCGGTGTACGTCGGGACAGGTGCGTGTTC 1970
Db 2184 ACTTGAGAGCTTGTATGGCATCAAGGACCAAGTTTATGTCCGAACTGGGTGTGTTTCA 2243
Qy 1971 GCGCCAGGCGCTCTACGGCTACAACTCCCAAGGAGCCCAAGA----- 2015
Db 2244 ACCGAACAGCTCTATATGTTGTTATGAGCCCCCAATTAAGCAGAAAGGGTGGTTTCTGT 2303
Qy 2016 -----GGCCCAAGATGTTGACCTGCGACTGCTGCCGTGCTTCGGCGCAAGA 2063
Db 2304 CATCACTATGTGGGGTAGGAGGAGGAGCAATCAAGAGGGCTCGGACAGAGAGA 2363
Qy 2064 AGCGGAACACGCCAAGAC-----GGGCTGCCGAGG 2096
Db 2364 AGTCGAGAGCATGTGACAGTCTCTGTGCCAGTATTCAACCTTTGAAGATATAGAGGAG 2423
Qy 2097 GCACCGCTGATATGGGAGTAGACACAGGAGATGCTCATGTCCCAATGAACTTCG 2156
Db 2424 GAGTTGAAGGGCTGGATTTGACGACGAGAAATCACCTTCTATGTCTCAATGAGCCTGG 2483
Qy 2157 AGAAGCGGTTCCGGCAGTCCGCGCGTTCGTACGTGCGAGCTGATGGAGGAAGGCGCG 2216
Db 2484 AGAAGATTTGGCCAGTCCGCGCGTGTGCTCCATCTGATGGAGTATGGTGGTG 2543
Qy 2217 TCCCTCTCGTCCAGCCCGCGCTCTCAAGAGGCGCATCCATGTATCAGTGGC 2276
Db 2544 TTCCTCAGTCGCAACTCCGAGTCTCTTCTGAAAGAAAGCTATCCATGTTTATAAGCTGTG 2603
Qy 2277 GCTACGAGGACAGACCGATCGGGGCTGGAGCTGGGGTGGATCTACGGGTCCGATCAGG 2336
Db 2604 GCTATAGGACAAAGACTGGAATGGGGAACCTGAGATCGGGTGGATCTACGGTTCGTGTGACG 2663
Qy 2337 AGGACATCTTACCGGGTTCAAGATGCACTGCGCGGGTGGCGTCCGTGTACTGCAATGC 2396
Db 2664 AAGACATCTCACCGGATTCAGATGACGCGCGAGGCTGGCGTCCGATCTACTGCAATGC 2723
Qy 2397 CGAAGCGGGCGGCTTCAAGGGGTCCGCGCGCGATCAATCTATCCGACCGTCTCAACCAGG 2456
Db 2724 CCAAGCGGCCAGCTTTCAAGGGGTCTGCCCCCATCAATCTTTCCGAGCGCTCTGAACCAGG 2783
Qy 2457 TGCTCCGGTGGCGCTGGGTCCGTCCAGATCTTCTTCAGCGGACAGGCCCTCGTGT 2516

Db 2784 TGCTCCGGTGGGCTCTTGGGTCCGTGGAGATCCTCTTTCAGCCGCGCACTGCCCTCTGGT 2843
Qy 2517 ACGGCTACAAAGAACGGCAACCTCAAGTGGCTGGAGCGCTTCGCTTACATCAACACACACCA 2576
Db 2844 ACGGCTAC--GGAGGGCGGCTCAAGTCTCTGGAGAGATTCGCGTACATCAACACACACCA 2900
Qy 2577 TCTAACCCCTTCACTCGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2636
Db 2901 TCTAACCCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2960
Qy 2637 TCACCGGCAAGTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2696
Db 2961 TCACCGGAAAGTTTCATATTCAGAGATCAGCAACTTCGCGCAGATCTGCTGCTGCTGCTGCT 3020
Qy 2697 TCTTCATGCTCACTTCGCGGACGCGGATCTGAGATGAGGTGAGGCGGGTGGAGCATCG 2756
Db 3021 TCTTCATCTGATCTTCGCGGACGCGGATCTGAGATGAGGTGAGGCGGGTGGAGCATCG 3080
Qy 2757 AGAGTGGTGGAGGAACGAGCAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2816
Db 3081 ACGAGTGGTGGAGGAACGAGCAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3140
Qy 2817 CGTCTGCTGAGGCGCTGCTCAAGTCTCTCGCGGATCGACACCACTTCACCGTCACT 2876
Db 3141 CGTGTTCAGGCGCTGCTCAAGTCTGCTGCGCGGATCGACACCACTTCACCGTCACT 3200
Qy 2877 CCAAGCGCACCGCGGACGAGGACGAGTTCGCGGATCTACGCTTCAAGTGGACCA 2936
Db 3201 CCAAGGCTCG---GACGAGGACGCGGACTTCGCGGAGCTGTACATGTTTCAAGTGGACGA 3257
Qy 2937 CGTCTCTATTCGCGGACCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2996
Db 3258 CGTCTCTGATTCGCGGACCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3317
Qy 2997 TCTCCGAGCGCATCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3056
Db 3318 TCTCTACGCGCATCAAGAGGATACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3377
Qy 3057 TCGCTCTTGGGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3116
Db 3378 TCGCTCTTGGGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3437
Qy 3117 ACAGGAGCGCCACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3176
Db 3438 ACCGCAACCGGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3497
Qy 3177 TCTGGGTGAGATCGACCCCTTTCATGCTGAGGACCAAGGCGCGGACGTCAGGCAAGTGTG 3236
Db 3498 TGTGGGTGCAATCGACCCCTTCACCAACCGCGCTCACTGCGCGCGGATACCAAGAGTGTG 3557
Qy 3237 GCATCAATTGCT 3248
Db 3558 GCATCAACTGCT 3569

RESULT 4

US-10-160-719A-25
; Sequence 25, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Heientjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthesases and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06

1394 CAGCCACCTTCTGCGAGCGCGCCGATGAAGAGAGATGATGAGAGTTCAAGTC 1453
1598 CAAACCTCATTTGTGAAGAACGCGCGCCATGAAGAGAGATGATGAAGATT 1657
1454 CGGATCAACGCGTGTGGCGCAAGCCCATGAAGTGCAGAGGAGGTGATCATGAAG 1513
1658 CGTATCAATGGTCTTTAGCCNAGGACAAAAGTTCCGAGGAGGATGATCATGCAA 1717
1514 GACGGCAAGCGGTGGCGGGAAACAAACCGCGCAACCCCGGCATGATCCAGGTTC 1573
1718 GATGGTACACCTTGGCGTGGAAACAATACATAGGAGACCATCTGGAATGATTCAGGTTTC 1777
1574 CTGGGCAAGCGCGCGCCACGACACGAGGCGAAGCTGCGCCGCTCGTACGTC 1633
1778 CTGGGTCAACAGTGGAGGCTTGAGCTTGAAGGCAATGAATCTCTGTTGGTTATGTG 1837
1634 TCCCGTGAAGACCGCGGATTCAGCACCAAGAAAGCGCGCCATGAAGCTCTG 1693
1838 TCTCGTGAAGACGCTCTGGATTCACATCACAAGAGGCTGTGCCATGATGCACTT 1897
1694 ATTGCGTCTCCGCGTGTGACCAACCGCGCATTCATGCTCAACTTGGACTGTGATCAC 1753
1898 GTTCGTGTATCAGTGTCTTACTAATGGGCAATACATGTTGAATCTTGAATGTGACCAC 1957
1754 TACATCAACACAGAGCCATCCGGAGGCGCATGTCTTCATGAGGACCTCAGGTC 1813
1958 TACATCAATAATAGCAAGGCTTCGAGAAAGCTATGTGCTTCTTATGAGACCAACCTA 2017
1814 GCGCGAAGGTCTGCTAGTTCAGTCCCGCAGAGTTTCAAGCATCGACGTCACGAC 1873
2018 GGAAGGAATGTCTGTATGTCCAATTTCTCAGAGGTTGATGTTATGAGGAATGAC 2077
1874 CGATACGCTAACAGGAACACCGTCTTCTCGACATCAACATGAAGGGCTGGAGCGATC 1933
2078 CGATATCAAAACAGGAACACTGTCTTTTCGATTAATTAACCTTGAGAGTCTTGACGGCAT 2137
1934 CAAGGCGGTGTACGTGCGGACAGGTTGGTTCGCGGCCAGGCGCTCTACGGTAC 1993
2138 CAAGGGCCAGTTATGTGGAACTGTGTGTGTTTAAACAGACGGCTTATATGTTAT 2197
1994 AACCTCCCAAGGACCAAGAGGCC-----CAAGATG 2026
2198 GAGCTTCAGTCAAGAAAAGAACGAGGCTTCTTCTCGCTTTGTGGGAGGAAA 2257
2027 GTGACTCGCATGCTGCCGTGCTCGCGCGCAAGAGCGGAACACGCCAAGGAC- -- 2083
2258 AAGACGTCAAAATCTAAGAGAGCTCGGAAAAGAAAGTCAATAGACAGCGACAGT 2317
2084 -----GGGCTGCGGAGGCGCCGCTGATATGGGAGTAGAT 2119
2318 TCTGTACCAAGTATTAATCTCGAAGATATAGAGGAAGGATGGAAGTTCTCAGTTTGTAT 2377
2120 AGCGACAAGGAGATGCTCATGTCCCAATGAATTCGAGAAGCGGTTTCGGGCGAGTCGCG 2179
2378 GATGAGAAATCGCTGATATGTTCTCAATGAGCTTGAGAGAGATTTGGCCAGTCCAGT 2437
2180 GCGTTCGTACGTGCGAGCTGATGGAGAGGCGCGTCCCTCTCTGTCGAGCGCCGCC 2239
2438 GTTTTTGTAGCTCTACTCTGTATGATATGGTGTCTTCACAACTCTGCAATCTCCAGAG 2497
2240 GCGCTCTCAAGAGGCGCATCATGTCATAGCTGCGGCTACGAGGACAGACCGACTGG 2299
2498 TCTCTTCTGAAGAAAGCTATTTCAATGATGCTGAGGATGAGGACAAAAGTCACTG 2557
2300 GGGCTGAGCTGGGTGATCTACGGTTCGATACGAGGAGCATCTTCAGCGGGTTCAAG 2359
2558 GGAAGTGAATTTGGTGGATCTATGGTTCTGTTACAGAGACATCTTCACCGGATTCAG 2617
2360 ATGCATGCGCGGGTGGCGTCTCGTGTACTGATGCGGAGGCGGGGTTCAAGGGG 2419
2618 ATGCATGCTCGAGGCTCGCATCAATCTACTGATGCTAAGCGACAGCTTCAAGGGA 2677
2420 TCGGCGCGGATCAATCTATCGGACCGCTCTCAACCAAGGTGCTCCGGTGGGCGCTG 2479

2678 TCTGCTCTATCAACCTTTCCGATCGTTTGAATCAAGTGTCTCGGTGGGCTCTTGGTTCC 2737
2480 GTGAGATCTTCTTACGCGGCGACAGCCCTGTGTACGGCTACAAAGAACGCAACCTC 2539
2738 ATTGAATCTTCTTACGAGGAGTGTCCCATATGATGGCT---ATGGAGGCGGCTT 2794
2540 AAGTGGCTGAGCGCTTTCGCTACATCAACACCAACATCTACCCCTTCACTCGCTCCG 2599
2795 AAATTCCTGAGAGATTTGCTTATCAACACAAACATTTATCCACTCATCAATCCG 2854
2600 CTGCTGCTTACTGCAACCTCCCGCGCTGCTCTCCACCGGCAAGTTTCATCATCGG 2659
2855 CTCCTCTGCTACTGATATTCGAGCAGTTTGTCTCTCACTGGGAAGTTTCATCATCCCA 2914
2660 TCGATTAGCAGTTTCCGACGCTTCTTCTCATGGCCCTTTCATGTCCATCTTCGCGACG 2719
2915 AAGATTAGTAACCTAGAGAGTGTGGTTTATATCGCTCTTATCTCAATCTTGGCCACT 2974
2720 GGCATCTCGAGATGCGGTGAGGCGGTGAGCATTCGAGAGTGTGGAGAAACGAGAG 2779
2975 GGTATCTCTGAGATGAGGTGGAGTGTGGCAATTCGATGAATGGTGGAGAAACGAGAG 3034
2780 TTCTGGGTCTATCGCGCGTGTCCGCGCATCTCTTCGCGCTCGTGCAGGCGCTCTCAAG 2839
3035 TTCTGGGTCTATGGTGTGATTTCTGGCGCATTTATTTGCGGCTTCTCCAGGGTCTCCTGAAG 3094
2840 GTCTCTCGCGGATCGACACAACTTCCACGCTCACTTCAAGGCGCACCGGCGACGAGGAC 2899
3095 GTGCTTGTGATCGACAGGCTTCACTGTCACTTCAAGGCGCAC---TGACGAAGAA 3151
2900 GACGAGTTCGCCAGCTCTACGCTTCAAGTGACCAACGCTCTCTCATCCGCGCCACGAG 2959
3152 GGTGATTTTCCGAGCTCTACATGTTCAAGTGAGCAACGCTTCTGATCCCAACCACT 3211
2960 CTGCTCATCATTAACGTCATCGCGCTGTCGCGGCGATCTCCGAGCGCATCAACAGGG 3019
3212 ATTTGATCATCAACCTGTCGCGGTGCTGCGCATTTCTACGCAATCAATAGCGGT 3271
3020 TACGAGTCTTGGGGGCGCTTCTCGGCAAGCTCTTCTTCGCTTCTGGGTCATCGTCCAC 3079
3272 TACGAGTCTGCGGAGCTCTTTTCGGAAGCTCTTCTTTCGCTTCTGGGTGATGTCCAC 3331
3080 CTCTACCGCTTCTCAAGGCGCTCATGGGCGCGCAAGAACAGGACCGCCACCGTTGTGTC 3139
3332 CTGTACCGCTTCTCAAGGCGCTCATGGGGAAGCAGAACCGCACCGCGCATTTGCTGTT 3391
3140 ATCTGGTCCATTTCTGTCGCTCCATCTTCTCCCTGCTCTGGGTGAGGATCGACCTTTC 3199
3392 GTCTGGGCTATCTCTCTTCTGCTGATTTCTTCTGATGTGGGTTCGTATCGATCCCATTC 3451
3200 ATGCTCAGGACCAAGGCGCGGAGCTCAGGAGTGTGGCATCAATTCCTGAGCTGTGTTAT 3259
3452 ACCACCGGCTCACTGGCCCTGATATCGGAATGGGATCAACTGCTAGGATGAGCTG 3511
3260 TAAGGTTCAAAATTCGAGCTTG 3283
3512 AAGATAGTTAAAGAGTGGAACTAG 3535

RESULT 8

US-10-160-719A-17
; Sequence 17, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helehtjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A


```
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3969
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (144)...(3401)
US-10-160-719A-17

Query Match      35.0%; Score 1204.6; DB 4; Length 3969;
Best Local Similarity 62.8%; Pred. No. 8e+214;
Matches 2051; Conservative 0; Mismatches 1099; Indels 117; Gaps 7;

QY 89 GCGATGGAGGCCAGCGCCGGCTGGTGGCGGCTCGCAACCGGAACGAGCTGGTGGCTG 148
DB 141 GCGATGGAGGCCAGCGCCGGCTGGTGGCGGCTCGCAACCGGAACGAGCTGGTGGCTG 200
QY 149 ATCCGGGGCCACAGAGACCC-----AAGCCGCTGCGGGCGGTGAGCGGGCAAGGTGTC 202
DB 201 ATCCGGCGGACCGCGGATCCCGGGCGGAAGCCCGCGGAGCAGAAACGCGGCAAGGTGTC 260
QY 203 GAGATATGGCGGACGAGCTCGGGCTCAGGTCGAGCGGCGACCTCTTCGTCGCTGCAAC 262
DB 261 CAGATTGCGGCGACGAGCTCGGCCCTTGGCCCCGGCGGGGACCCCTTCGTGGCGTGCAC 320
QY 263 GAGTGGCGGCTTCCCGCTGTGCGCGCCCTGCTACGAGTACGAGCGCGGGAGGGCAAGCAG 322
DB 321 GAGTGGCGCTTCCCGCTGTGCGCGGACTGCTACGAGTACGAGCGCGGGAGGGCAAGCAG 380
QY 323 AACTGCCCCAGTGCAGAGCGCTCAAGCGCTCAAGGGGAGCGCGAGGGTTCGCGG 382
DB 381 AACTGCCCCAGTGCAGAGCTCGATACAGCGCTCAAGGGGCTCAAGGGGCTCAAGGGTGC 440
QY 383 GAGGATGAGGAGGAGACATCGAGACCTGGAGCAGGTTCAACATCGACGAGAGAT 442
DB 441 GAGGAGGAGGAGCGCGCTGATGACCTGGACCAACGAGTTCACATGGGACGCGCATGAC 500
QY 443 CAGCAGAGGCGAGCTGGAGGSCA-----ACATGCAAGAACAGCCAGATCACCGAGGCGATG 496
DB 501 TCGCAGTCTGTGCCGATCCATGCTCTACGGCCACATGAGCTACGCGCGTGGAGGTGAC 560
QY 497 CTGCACGGCAGGATGAGCTACGGGAGGGGCGCCGAGCAGCGGCGAACAAACCCCG 556
DB 561 CCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTCACCTCTCCACCAACGGG 620
QY 557 CAGATCCCGCCCATCATCACCAGCTCCGCTCCGCTCGCTGCGGTGAGCGGTTCGATTT 616
DB 621 CAAATGGTGGATGACATCCACCGGAGCAGCAGCGCTGGTGGCTTCTTTCATGGGTGGT 680
QY 617 ACCAAGCGGTATGGCCACGGCGAGGCTCTGCTCTTCCTTGCAACGCGCATCCATCCGTAC 676
DB 681 GGGGGAAGAGGATACATCCCTCTCTTATGCGGATCCCGACTTACTGTGCNACCCAGG 740
QY 677 CTGTGTCTGAGCCAGGGAGTGCAGAGTGGGACGAGAGAAAGAGTGAAGTGAAGGAG 736
DB 741 TCTATGGACCCATCCCAAGGATCTTGTGTCATATGGGTATGGTGTGCTTGGAGGAA 800
QY 737 AGGATGAGACGACTGGAGTCCAGCAGGGGATCTCTCGC-----GGCGGCGCC 784
DB 801 CGGATGGAGAAATGGGAAGCAGACAGAGAGAGATGACACAGCGGGGAATGATGGTGGT 860
QY 785 GATCCCGAAGACATGAGCGCGGAGTGGCACTCAACGAGCGAGGCGGAGCGCGCTGTG 844
DB 861 GGTGATGATGGTACGATGCTGATCTACCACTAATGATGATGATGATGATGATGATGATG 920
QY 845 AGGAAGGTGTCGATCGCGTCGAGCAAGGTGAACCCCGTACCGGATGATGATGATGATGATG 904
```

```
DB 921 AGGAAAAATCCACTTCCACTCAAGCCAGATTAATCCATATAGGATGATTAATCATTTTCG 980
QY 905 CTCGTTGTGCTCGCTCTCTCTCGGTACCGTATCTCTCGACCCCGTCCCGGACGCATC 964
DB 981 CTTGTGTTTGGGGTCTCTCTCACACCGAGTATGATCCCGTGAATGATGATTT 1040
QY 965 GGGCTGTGGCTCTCTCCATCATCTCGAGATCTGTTGCGCATCTCTCTGATCTCTCGAC 1024
DB 1041 GCTTTGTGGCTCATATCTGTTATCTGTAATCTGTTTGGCATGCTTGGATTTCTGAT 1100
QY 1025 CAGTTCCCAAGTGGTTCCTCCATCGACCGCGAGACGTACTCGACCGCTCTCCCTCAG 1084
DB 1101 CAAATCCCAAGTGGTTCCTTATGAGAGAGAGACTTACTAGACCGGCTGTCACTGAG 1160
QY 1085 TACGAGAGGGAAGGGAGCGCTGCTGCTGCGGGTGGACCTGTTCTGTGAGCAGCGTG 1144
DB 1161 TTCGACAGGAAGGCCAGCATCTCAACTTGTCTCAATGATTTCTTTGTGATGATGAT 1220
QY 1145 GACCGCTCAAGGAGCGCGCTGCTGACCGCAACACACCGTCTCTCTCACTCTCGCGTA 1204
DB 1221 GATCCCTTAAAGAACTCTCTTTGGTCAACAATACTGTTCTATCTATCTCTTTCGGTG 1280
QY 1205 GACTACCCCGTGGACAAGGTCTCTGCTACGTCTCGACGACGGCGGCTGATGCTGACG 1264
DB 1281 GATTATCTCTGTTGATAAGGTTCTTGTCTATGTTCTGATGATGCTGCTCAATGCTAAC 1340
QY 1265 TTCGAGTCTGCTCGGAGACGGCGGTTGCGCGCAAGTGGTGGCTCTCTGCAAGAAG 1324
DB 1341 TTTGAAGCAATATCTGAAACATCTGAATTTGAAAGAAATGGGTCTCTTTCTGCAACGG 1400
QY 1325 TTCGCGATCGAGCCCGCGCGCTGCTCTCTCTCAAGTCTCAAGTCTCAAGTCTCAAG 1384
DB 1401 TACAATATTGAACCTCCGCTCCAGAGTGTACTTCCACAGAGATAGACTACTTTGAAA 1460
QY 1385 GACAGAGTGCAGCCACCTCTGTCGAGAGCGCGGCGCATGAAGAGAGATGATGAGGAG 1444
DB 1461 GACAGGTGCGACGCAAACTTTGTTAGGAGAGAGAGCAATGAAGAGAGATGATGAGGAA 1520
QY 1445 TTCAGGTCCCGATCAACGCGCTGTTGGCCAAAGGCTGAAAGTGCAGGAGGGGTG 1504
DB 1521 TTCAGGTGAGAAATCAATGCTTGTGTCAAAGCCAGAAAGTTCCTGAAGAGATGG 1580
QY 1505 ATCAAGAGGACGACGCGCTGCGCGCGGAAACAACACCGCGACCAACCGCGCATGATC 1564
DB 1581 ACAATGCAAGATGGAACCCCTGGCTGGAACAATGTTGCTGATCATCTGGAATGAT 1640
QY 1565 CAGGTGTTCTCGGCGCACAGCGCGCGCGCACGACACGAGGGGCAACGAGCTGCGCGCTC 1624
DB 1641 CAGGTCTTCTTGGCCAAAGCGGAGGCTTGAAGTGTGAGGGGAAATGAACCTGCCACG 1700
QY 1625 GTGTAGTCTCCGTTGAGAGCGCGCGGATTCAGACCAACAAGAGCGCGCGCATG 1684
DB 1701 GTTTATGTTCTAGAGAGAAACGACAGGCTATACCATCATGAAGAACTGGTGTATG 1760
QY 1685 AACGCTCTGATTTGGGCTCTCCCGCGTGTGACCAACGCGCCATTCATGCTCAACTTGGAC 1744
DB 1761 AATGATTTGGTCCGAGTCTCTGCTGACTAACAAATGCTCCATATTTGTTAACTTGGAT 1820
QY 1745 TGTGATCACTACATCAACAGAGAGGCGCATCCGGAGGCGCATGCTCTCTCATGGAC 1804
DB 1821 TGTGATCACTACATCAACAGAGGCTATGAAGAGCAATGTTTATGATGATGATGATG 1880
QY 1805 CCTCAGGTCCGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1864
DB 1881 CTTTATCTAGGAAGAGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1940
QY 1865 GTGCACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1924
DB 1941 CGCCATGACCGATATGCTAAACCGGAATGTTGCTTTTGTATATCAACATGAAGGTTG 2000
QY 1925 GACGGATCCAAAGGCGCGGTGATGCTGCGGACAGGCTGCTGCTGCGGCGCGGCGCTC 1984
```


Db 2001 GATGGTATTGAGGTCCAAATTTATGTTGGTACTGATGTGTATTTAGAGCGAGGCATT 2060
Qy 1985 TAGGGTACAAACCCCTCCCAAGGGACCAAGAGGCC-----CAAGATG 2026
Db 2061 TATGGTTATGATGCCCCCAAAACAAAGAGCCACCATCAAGGACTTGCAACTGCTGCC 2120
Qy 2027 GTGACCTGCGACTGCTGCCGCTCTCCGCGCGCAAGAGCGGAACACGCAAGGACGG 2086
Db 2121 AAGTGGTCTTTTGTCTGCTTTGCGCAATAGGAAGCAAAAGAAAGACTTACCAAAACC 2180
Qy 2087 CTCGCCGAG----- 2095
Db 2181 AAAACAGAGAAAGAAATTTATTTTCAAGAAAGAGAACCAATCCCTGTCATAT 2240
Qy 2096 -----GGCACCGCTGATATGGAGTAGATAGCGCAAGAGAGATGCTCATG 2140
Db 2241 GCTCTTGGTGAATTTGAGAGAGCTGCTCCAGGAGCTGAGAAATGAAAGGCGGTATTGTA 2300
Qy 2141 TCCACATGAACTTCGAGAGAGCGTTCCGGCAGTCCCGCGGTTTCGTCAGTCCAGCTG 2200
Db 2301 AATCAACAAAAATTTAGAAAGAAATTTGGCCAAATCTTCTGTTTGTATCATCCACACT 2360
Qy 2201 ATGAGGAAGCGCGCTCCCTCTCTCGTCCAGCCCGCGCTCTCAAGGAGGCCATC 2260
Db 2361 CTCGAGATGGTGAACCTTGAAGAGTGCAGTCTCTCTTTTGAAGAGCTATA 2420
Qy 2261 CATGTCTATAGCTCGCGCTACGAGCAAGACCGACTGGGGCTGGAGCTGGGGTGGATC 2320
Db 2421 CATGTCTATTAGTTGCTGTTATGAAGACAGACAGACTGGGGAAGAGATTGGCTGGATC 2480
Qy 2321 TAGGGTGCATACGAGGAGCATCTGACCGGGTTCAAGATGCACTCCCGGGTGGCGC 2380
Db 2481 TATGGATCAGTTACAGAGATATTCTAACTGGTTTCAAGATGCAATGCTATGTTGGCGG 2540
Qy 2381 TCCGTGTACTGCTACCGAGCGGGCGCTTCAAGGGGTGGCGCGATCAATCTATCG 2440
Db 2541 TCAATTTACTGATACCTAAACGGGTGCAATCAAGGTTCTGCACCTCTGAATCTTTCA 2600
Qy 2441 GACCGTCTCAACAGAGTGTCCGGTGGCGCTCGGGTCCGTTCGAGATCTTTTCAGCGG 2500
Db 2601 GATCGTCTTCCACAGGTGCTCGGTGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAAT 2660
Qy 2501 CACAGCCCCGTGTAGGCTACAGAAACCGCACTCAAGTGGCTGGAGCGCTTCGCC 2560
Db 2661 CATTGCCCTCTTGGTATGGGT---ATGGTGGCGGTCTGAAATTTTGGAAAGATTTTC 2717
Qy 2561 TACATCAACACCACTACCTCTCACCTCGCTCCCGCTGCTCGCTACTGCAACCTC 2620
Db 2718 TACATCACTCATCTGCTGATCTCTTGAGCAATCTATCCCTCTTGGCTTACTGTACATTG 2777
Qy 2621 CCCCGGTCTGCTCTCACCGCAAGTTTCATATGCGCTCGATAGCAAGTTTCGCCAGC 2680
Db 2778 CCGTCCATCTGTTTATTGACAGGGAATTTATCACTCCAGAGCTGAAATATGTTGCCAGC 2837
Qy 2681 CTCTTCTTATCGCCCTCTTCATGTCATCTTCCGAGCGGGCATCTCGAGAGATCGCGTGG 2740
Db 2838 CTGTGGTTTATGTCACCTTTTATCTGCAATTTTGTCTACGAGCATCTCTAGAAATGAGATG 2897
Qy 2741 AGCGGGTGAGCATCGAGGAGTGTGGAGGAACGAGCAGTTCTGGGTCTATCGCGCGCTG 2800
Db 2898 AGTGGTGTGGAATGTATGACTGTGGAGGAATGAGCAGTTCTGGGTCTATGAGAGTGTG 2957
Qy 2801 TCCGCGCATCTCTTCCGCTGTCGAGGGGCTGCTCAAGGTCTCTCGCGGGATCGACACC 2860
Db 2958 TCCTCACACCTCTTGTGTGTTCCAGGGACTTCTCAAGGTCTATAGTGGTGTGATACA 3017
Qy 2861 AACTTCACCTGCTACCTCAAGGCCACCGGCGAGAGAGACAGAGTTTCGCCGAGCTCTAC 2920
Db 3018 AGCTTCACCTGATCAAAAG-----GGTGGAGATGATGAGGAGTTCTCAGAGCTATAT 3071
Qy 2921 GCCTTCAAGTGGACCACTCTCATCCCGCCACACCGTCTCATCATTTACGTCATC 2980
Db 3072 ACATTCAAATGGACTACCTTATTGATACCTCTTACCACCTTGTCTTCTATTGAACCTTCATT 3131

Qy 2981 GGCGTGTGGCCCGCATCTCCGACCGCATCAACAAACGGGTACCAAGTCTCTGGGGGCCCTC 3040
Db 3132 GGTGTGTGCTGCTGGCGTTTCAAAATGCGATCAATAACGGATATGAGTCAATGGGGCCCTC 3191
Qy 3041 TTCGGCAAGCTCTTCTTCCGCTTCTGGGTTCATGCTCCACCTCTACCCGTTCTCAAGGG 3100
Db 3192 TTTGGGAGCTATCTTGTGCAATTTGGTGATGTGTCATCTTTATCCCTTTCTCAAAAGT 3251
Qy 3101 CTCATGGGGCGCAGAACAGAGCGCCACCGTTTGTGTATCTGTGTCATTTCTGTGCCC 3160
Db 3252 TTGTTTGAAGGCAAAACAGACACCAACGATTGTTCATGCTCTGTTGTCATTCTGTGCT 3311
Qy 3161 TCCATCTTCTCCTGCTCTGGGTTCAGATGCAACCTTTTCATGCTCAGGACCAAGGGCCG 3220
Db 3312 TCAATCTTCTGCTCTCTTGGGTTCGGAATGATCTTCTTCCGAGGATGATGTTGCTCG 3371
Qy 3221 GACGTCCAGGAGTGTGSCATCAATTGC 3247
Db 3372 CTCTTGGAGAGTGTGTTGGATTGC 3398

RESULT 9

US-10-160-719A-37
; Sequence 37, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 3969
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (144)...(3401)
US-10-160-719A-37

Query Match 35.0%; Score 1204.6; DB 4; Length 3969;
Best Local Similarity 62.8%; Pred. No. 8e-214;
Matches 2051; Conservative 0; Mismatches 1099; Indels 117; Gaps 7;

Qy 89 GCATGAGAGCCAGCCCGCGCTGTGGCGCTGCGACCAACCGGAAACAGCTGTGGTGTG 148
Db 141 GCGATGAGAGCGAGCGCGCGCTGTGGCGCTGCCACACCGGCAACGAGCTCGTGTG 200
Qy 149 ATCCGGGGCCACAGAGACCCC-----AAGCCGCTGGGGCGCTGAGCGGCGAGTGTG 202
Db 201 ATCCGGCGGACGGCGATCCCGGGCGGAAGCCCGCGGAGCAGAACCGGCGAGTGTG 260
Qy 203 GAGATATGGCGGCAACAGAGTCCGGCTCAGCGTGGAGCGGCACTCTTTCGCTGCAAC 262
Db 261 CAGATTTGGCGCAACGAGTCCGGCTTGGCCCCGGGGGACCCCTTCGTTGGCTGCAAC 320
Qy 263 GAGTGGCGTTCCCGTGTGGCGCCCTGCTACGATACGAGCGCGGAGGAGGACGAG 322
Db 321 GAGTGGCGTTCCCGTGTGGCGGACTGCTACGAATACGAGCGCGGAGGAGGACGAG 380
Qy 323 AACTGCCCCAGTGTGAAGAGCGGCTCAAGCGCTCAAGGGAGCCCGAGGTTGCCGG 382

Db 381 AACTGCCCCAGTGCAGACTCGATACAGAGCGCTCAAGGGCTGCCAAGCTGTGACCGGT 440
QY 383 GACGATGACGAGGAGGACATCGACGACCTTGGAGCAGAGTTCAACATCGACGACGAGAAT 442
Db 441 GACGAGGAGGAGGACGGCGTCGATGACCTTGGACAAAGAGTTCACTGGGACGGCCATGAC 500
QY 443 CACGAGGAGGAGTGGAGGGCA-----ACATCGAGAAAGCCAGATGACCGGAGGCGATG 496
Db 501 TCGCAGTCTGTGCGCGAGTCCATGCTCTACGGCCACATGAGCTACGGCGCTGGAGGTGAC 560
QY 497 CTCGACGGCAGGATGAGTACGGGAGGGGCCGACGACGGCGACGGCAACACCCCG 556
Db 561 CCTAATGGCCGCCCAAGCTTTCCAGCTCAACCCCAATGTTCCACTCTCTCAACACGGG 620
QY 557 CAGATCCCGCCCATCATCACCGGCTCCCGCTCCGGTGGAGCGGTGAGTTCCGAT 616
Db 621 CAAATGGTGATGACATCCACCGGAGCAGCAGCGCTGGTCTTTTCATGGGTGGT 680
QY 617 ACCAAGGGTATGGCCACGGCGAGGTCTCGTCTTCCCTGTGCAAGCGCATCCATCCGTAC 676
Db 681 GGGGAAAGAGGATACATCCCTTCTTATGCGGATCCAGCTTACCTGTGCAACCCAGG 740
QY 677 CCTGTGCTCGAGCCAGGAGTGCAGTGGGACGAGAGAAAGAGTGAAGTGGAGGAG 736
Db 741 TCTATGACCCCATCCAAAGATCTTGTGCTATGCGGTATGAGTGTGCTTGGAAAGAA 800
QY 737 AGGATGACGACTGGAGTCCAAAGCAGGCGATCCTCGGC-----GGCGCGCC 784
Db 801 CGGATGGAGAAATTGGAAGCAGAGACAAAGAGAGATGCACCAGAGGGGAATGATGTGT 860
QY 785 GATCCCGAAGACATGAGCGCGAGCTGGGCACTGAAAGCAGAGCGGAGGACGGCTGTGCG 844
Db 861 GGTGATGATGGTGACGATGCTGATCTACCATAATGGAATGAGTGAAGCAACAACCTGTCC 920
QY 845 AGGAAGGTGCGATCGCTCGACGAGGTGAACCCGTACCGGATGCTGATCGTGTGCGT 904
Db 921 AGGAAATTCACCTTCCATCAAGCCAGATTAATCCATATAGGATGATTAATCATATTGCG 980
QY 905 CTGCTGTGTGCTCGCTTTCTTCTCCGATACCGTATCTGCACCCCGTCCCGGACGCCATC 964
Db 981 CTGTGTGTTTTGGGTTCTTCTTCCACTACCGAGTATGATCCGGTGAATGATCATTT 1040
QY 965 GGGCTGTGCTGTCTCATCATCTCGGAGATCTGGTTGCGCATCTCTGATGATCTCGAC 1024
Db 1041 GCTTTGTGGCTCATCTGTATCTGTGAATCTGGTTTGCCATGTCTTGGATCTTGTAT 1100
QY 1025 CAGTTCCTCCAAAGTGTTCCTCATCGACCGCGAGACGTACCTCGACCGCTCTCCCTCAG 1084
Db 1101 CAATTCCTCAAGTGTTCCTTATGAGAGAGACTTACCCTAGACCGCTGTCACTGAGG 1160
QY 1085 TACGAGAGGAGGAGCGCTGCTGCTCGCGGTGAGCCTGTTCTGTGAGCAGCGTG 1144
Db 1161 TTCGACAAAGAGCCAGCCATCTCAACTGTGCTCCAAATGATTTCTTTGTGACGAGTT 1220
QY 1145 GACCCGCTCAAGAGCGCGCTGGTGAACCGCAACACCGTGTCTCTCAATCTCTCGCGGTA 1204
Db 1221 GATCCCTTAAAGGAACTCTCTTGGTCACAACAATACTGTTCTATCTATCTCTTTCGGTG 1280
QY 1205 GACTACCCGTGACAGGTCTCTGTAGTCTCCGACGACCGCGCTCGATGCTGACG 1264
Db 1281 GATTATCTGTTGATAGGTTTCTTGTATGTTCTGTATGATGGTGTGCAATGCTAAACG 1340
QY 1265 TTCAGTCTGTCTGGAGACGGCGGATTCGCGCGCAAGTGGGTGCGCTTCTTCAAGAAG 1324
Db 1341 TTTGAAGCATTAATCTGAACATCTGAATTTGCAAGAAATGGGTTCTTTCTGCAACCG 1400
QY 1325 TTCGGATCGAGCCCGCGCGGAGTTCTACTTCTGTGCTCAAGTTCGACTACCTCAAG 1384
Db 1401 TACAATAATTGAACCTCGCGCTCCAGAGTGTACTTCCAAAGAGATAGACTACTTTGAAA 1460
QY 1385 GACAAGGTGACGCCCACTTCTGTGAGGAGCGCGCCCATGAAGAGAGATGATAGGAG 1444
Db 1461 GACAAGGTGACCAAACTTTGTAGGAGAGGAGAGCAATGAAGAGAGTATAGGAA 1520

QY 1445 TTCAAGGTCGGATCAACCGCTGTGGCCAAAGGCOATGAAGTGCCTGGCAGAGGGGTG 1504
Db 1521 TTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCAGAAAGTTCTCTGAAGAAGGATG 1580
QY 1505 ATCATGAAGGACGGCAGCGCTGGCCCGGGAACAACACCGCGACACCCCGCATGATC 1564
Db 1581 ACAATGCAAGATGGAAACCCCTGGCTGGAAAACAATGTTCTGATCATCTCGGAATGAT 1640
QY 1565 CAGGTGTTCTCGGCCACAGCGCGCACGACACCGAGGCAACGAGCTGCCCGCCCTC 1624
Db 1641 CAGGTCTTCTTGGCCAAAGCGGAGGCTTGAATGTGAGGAAATGAATGCCACGATTG 1700
QY 1625 GTGTACGTCTCCCGTGAGAGCGCCCGGATTTCCAGCACCAAGAAGCGCGCGCATG 1684
Db 1701 GTTTATGTTTCTAGAGAGAAACGACCGAGCTATAACCATATAAGAAGCTGGTGCTATG 1760
QY 1685 AAGCTCTGATTCGGCTCTCCCGCTGTGACCAACCGCCCATTTCAATGCTCAATCTGGAC 1744
Db 1761 AATGCAATTTGGTCCGAGTCTCTCTGTACTAACAAATGCTCCATATTTGTTAACTTTGGAT 1820
QY 1745 TGTGATCACTACATCAACAACAGCAGGCGCATCCGGGAGCCATGTCTTCTCATGAGC 1804
Db 1821 TGTGATCACTACATCAACAACAGCAGGCTATAAGGAAGCAATGTGTTTATGATGGAC 1880
QY 1805 CCTCAGGTCGGCGGAAGGTCTGTACGTTCTAGTTTCCGCGAGAGTTTCGACGSCATCGAC 1864
Db 1881 CCTTTACTAGGAAAGAGGTTTGTATGTACAGTTCCCTCAAAGATTGATGGATTGAT 1940
QY 1865 GTGACGACCGATAGCTTAAAGGAACACCGTCTTCTCGACATCAACATGAAGGGCTG 1924
Db 1941 CGCCATGACCGATATGTCTAACCGGAATGTGTCTTTTGTATCAACATGAAGGTTT 2000
QY 1925 GACGGCATCAAGGCCGGTGTAGCTCGGACAGGCTGTGTTCCGGCGCCAGGCGCTC 1984
Db 2001 GATGGTATTCAGGTTCCAAATTTATGTTGGTACTGATGTGTTATTAGAAGCGAGGCATTA 2060
QY 1985 TACGGCTACAACCTCCCAAGGGAACCAAGAGGCC-----CAAGATG 2026
Db 2061 TATGGTTATGATGCCCCCAAAACAAAGAGCCACCATTCAAGGACTTGTCAACTGCTGCC 2120
QY 2027 GTGACCTGGCATGTCTGCCCGCTGTCTTCCGCGGAGAGAGCGGAACACACCGAAGGCGG 2086
Db 2121 AAGTGGTGTCTTTTGTCTGTGCTTGTGGCAATAGGAAGCAAAAGAAAGACTACCAACCC 2180
QY 2087 CTGCCGAG----- 2095
Db 2181 AAAACAGAGAAAGAAAGTTATATTTTCAAGAAAGAGAAACCAATCCCCTGCATAT 2240
QY 2096 -----GGCACCGCTGATATGGAGTAGATAGCGAACAGGAGATGCTCATG 2140
Db 2241 GCTCTTGGTGAATTTGACGAAGCTGTCTCCAGGAGCTGAGAAATGAAGGCGCGGTAATGTA 2300
QY 2141 TCCCATCATGAATCTCGAAGAGCGTTCCGGGAGTCCGGCGGCTTGTGTGACGTGACGCTG 2200
Db 2301 AATCAACAAAAATTAGAAAGAAATTTGGCCAAATCTTCTGTTTGTGTTATCATCCACACT 2360
QY 2201 ATGGAGGAGGCGGCTCCTCTCTGTCGAGCCCGCGCTCAAGGAGGCGCATC 2260
Db 2361 CTCGAGAAATGGTGGAAACCTTGAAGAGTGAAGTCTGCTCTCTTTTGAAGAGCTATA 2420
QY 2261 CATGTCAATCAGCTCGGCTACGAGGAACAAGCCGACTGGGGCTGGAGCTGGGGTGGATC 2320
Db 2421 CATGTCAATTTGTTGTTATGAAGACAGACAGCTGGGGAAGAGATGGCTGGATC 2480
QY 2321 TAGGGTTCATCAAGGAGCATCTGACGGGTTCAAGATGCACTGCCCGGGTGGCGC 2380
Db 2481 TATGGATCAGTTACGAAGATATTCTAACTGGTTTCAAGATGCAATTTGTCTATGTTGGCGG 2540
QY 2381 TCCGTGTACTGATGCCGAAGCGGGGCTTCAAGGGCTCGCGCGGATCAATCTATCTG 2440
Db 2541 TCAATTTACTGTACCTTAAACGGGTTGCAATTCMAAGGTTCTGCACCTCTGATCTTTCA 2600

Qy 2441 GACCGTCTCAACAGGCTCGGTGGGGTGGGGTCCGTGAGATCTTCTTCCAGCGG 2500
Db 2601 GATCGTCTTCAACAGGCTCTGGGGTGGGGTCTTGGGTCTATTGAGATCTTCTTCCAGCAAT 2660
Qy 2501 CACAGCCCTCTGCTGATCGGCTCAAGAAACGCAACCTCAAGTGGCTGGAGCGTTCGCC 2560
Db 2661 CATTGCCCTCTTGGTATGGGT---ATGGTGGCGGTCTGAATTTTGGAAAGATTTC 2717
Qy 2561 TACATCAACACACCATCTACCCCTTACCTCGGTCCGGCTGCTGCGCTACTGCAACCTC 2620
Db 2718 TACATCAACCTCATCGTATCTTGGACATCTATCCCTCTTGGCTTACTGTACATG 2777
Qy 2621 CCCCGCGTCTGCTCTCACCGGCAAGTTCTATCTGCGTCTGATTTAGACCTTCCGCGC 2680
Db 2778 CTTGCCATCTGTTATTGACAGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGC 2837
Qy 2681 CTCTTCTTTCATCGCCCTCTTCAATGTCATCTTTCGCGAGCGGCATCTTGGAGATCGGTGG 2740
Db 2838 CTGTGGTTTCATGTCATCTTTTATCTGCAATTTTGTACGAGCATCTTAGAATGAGATGG 2897
Qy 2741 AGCGGGTGGATCGAGGAGTGTGGAGGAACGAGCAGTGTCTGGGTCTATCGCGGGGTG 2800
Db 2898 AGTGGTGTGGAATTTGATGACTGTGGTGGAGGAATGAGCAGTCTTGGGTCAITGGAGGTG 2957
Qy 2801 TCCGGCATCTTTCGCGCTGCTGAGGCGCTCTCAAGTCTCTGCGCGGATCGACAC 2860
Db 2958 TCCTCACACCTCTTGTGCTGTTCCAGGACTTCTCAAGGTCTAGCTGTGTTGATACA 3017
Qy 2861 AACTTCAACCGTCACCTCAAGGCGCACCGCGAGGAGGACGAGTTCGCGGAGCTCTAC 2920
Db 3018 AGCTTCAACCGTGACATCAAG-----GTTGGAGATGATGAGGAGTCTCAGAGCTATAT 3071
Qy 2921 GCCTTCAAGTGGACACGCTCTCATCCCGCCCAACCGCTGCTCATCAATTAAGCTCATC 2980
Db 3072 ACATTCAAATGGACTACCTTATGATACCTCTACACCTTGTCTTATTGAACTTCATT 3131
Qy 2981 GGCTCTGTCGGCGCATCTCGAGCGCCATCAACACGGGTACAGTCTTGGGGGCCCCC 3040
Db 3132 GGTGTGTCGTCGGCGTTTCAATGCGATCAATAACGGATGATGATCATGGGGCCCCC 3191
Qy 3041 TTCGCGAAGCTCTTCTTCCGCTCTTGGGTCTATCGTCAACCTCTTACCGGTTCTCAAGGG 3100
Db 3192 TTTGGGAGCTATCTTGTGCAATTTGGGTGATTTGCACTTTATCCCTTCTCAAGGT 3251
Qy 3101 CTATGGGGCCCAAGAACAGACGCCACCGTTGTGTCATCTGCTCATCTCTGCGCC 3160
Db 3252 TTGGTTGGAAGGCAAAACAGGACCAACAGATTGTCTGCTGCTGCTCATCTGCTGCT 3311
Qy 3161 TCCATCTTCTCCCTGCTGCTGGTCAAGTCAACCGCTTTCATCTGCTCAGGACCAAGGGCCG 3220
Db 3312 TCAATCTTCTCGCTCTTGGGTTCGGATTGATCTTCTTTCGGAAGGATGATGGTCCG 3371
Qy 3221 GACGTCAAGGAGTGTGGCATCAATTGC 3247
Db 3372 CTTCTTGGAGAGTGTGTTGGATTGC 3398

RESULT 10

US-10-160-719A-13
; Sequence 13, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17

; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 3725
; TYPE: DNA
; ORGANISM: Zea mays
; NAME/KEY: CDS
; LOCATION: (179) ... (3400)
US-10-160-719A-13

Query Match 34.5%; Score 1187.2; DB 4; Length 3725;
Best Local Similarity 62.8%; Pred. No. 1.3e-210;
Matches 2033; Conservative 0; Mismatches 1123; Indels 82; Gaps 9;

Qy 89 GCGATGAGGCGCAGCGCGCGGTGGTGGCGGCTCGCACAAACCGGAACGAGCTGCTGCTG 148
Db 176 GAGATGCGGCGCAACAAGGGGATGGTGGCAGGCTCTCACAAACCGCAACGAGTTCGTCTG 235
Qy 149 ATCCGGGGCCACGAGGACCC-----AAGCCGCTGCGGGCGCTGAGCGGGCAG 196
Db 236 ATCCGCGCACGACGCGCGCGCTGTCCCGCTAAGCCACGAAGAGTGCGAATGGGCG 295
Qy 197 GTGTGCGAGATATGCGCGGAGCGAGGTGCGGCTCAGGTGGACGCGGACCTCTTCTGTCGCC 256
Db 296 GTCTGCGCAGATTTGTGGCGACACTGTTGGGGTTTCAGCCACTGCTGATGTTGTTGCC 355
Qy 257 TGCACACAGTGGCGCTTCCCGTGTGCGCGGCTGTACGAGTACGAGCGCGCGGAGGGC 316
Db 356 TGCATGAGTGTGCTTCTTCTGCTGCGCGCTTGTATGAGTACGAGCGCAAGAGGG 415
Qy 317 ACGCAGAACTGCCCCCAGTGCAGACGCGCTCAAGCGCCTCAAGGGGAGCCCGAGGGTT 376
Db 416 AACCAATGCTGCTCAGTGCAGACTAGATACAGAGACAGAAAGGTAGCCCTCGAGTT 475
Qy 377 GCGCGGACGATGACGAGGAGGACATCGACGACTGTGAGACGAGTTCACATCGACAC 436
Db 476 CATGCTGATGATGAGGAGGAGATGTTGATGACCTCGGCAATGAATTCACATATAAGCAA 535
Qy 437 GAGATTCAGCAGAGGCGAGCTGGAGGCAACATGACAGAACCCAGATCACCGAGGCGATG 496
Db 536 GGCATGGGAAGGGCC--CAGAGTGGCAGCTTCAAGAGATGACGCTGATCTGTCTTCAT 593
Qy 497 CTGCAACGCGAGGATGAGCTACGAGGAGGGGCCCGACGACGCGCGCAGCAACACCCCG 556
Db 594 CTGCTCGCATGACCCACACCATCGGATTCCAGGCTTACAGGCTTACAGTGGACACAGATATCG 653
Qy 557 CAGATCCCGCC---CATCATACCGGCTCCCGTCCGTCGCGGTGAGCGGTGAGTTCC 612
Db 654 GAGAGATCTCTGATGATCTCCCTGACCGCTCATTTCTATCCGAGTCCCAACATCGAGCTATG 713
Qy 613 GATTACCAACGGGTATGCGCCACGCGAGGCTCGTCTTCCCTGACAAAGCGCATCCATCC 672
Db 714 TTGATCCAAAGGTTTCCAGTTCTCTGTGAGGANTGTGGACCCCTCGAAGACTTGAATTCCT 773
Qy 673 GTACCCCTGTCTGAGCGAGGAGTGCCAAAGTGGGAGAGAAAGAAAGTGAAGCTG-GA 731
Db 774 ATGGGCTTAATAGTGTGACTGGNAGGAAGAGTTGAGAGCTGGAGGTTAAACAGGACA 833
Qy 732 AGGAGAGGATGAGAGCTGGAAGTCCAAAGCAGGAGCATCTCTCGCGCGCGCGGCGGATCCCG 791
Db 834 AAAATATGTTGCAAGTGAATTAATAATATCCAGAGGCTAGAGGAGACATGAGAGGGGACTG 893
Qy 792 AAGACATGGACCGCGAGCTGCGCACTGAAACGAGGCGAGGAGCGCGCTGTCTGAGGAGG 851
Db 894 GCTCAATGGAGAGATATGCAATGTTGATGATGACCGCTTACCTTTGAGCCGCAATTG 953
Qy 852 TGTGATCGGCTCGAGCAAGTGAACCCGCTACCGGATGTTGATCGTGGTGGCTCTCGTTG 911
Db 954 TGCCAAATTTCTCAACACAGCTCAACCTTTTACCGGATAGTAATCATCTTCCGCTTATCA 1013

Db	3168	GCTACCAATCCTGGGGTCCGCTCTTTGGAAAGCTGTCTTCTCGATCTGGGTGATCTCC	3227
Qy	3078	ACCTCTACCCGTTCCCTCAAGGGGCTCATGGGGGCCAGAAACAGAGACGCCACCGCTGTGTG	3137
Db	3228	ATCTTACCCCTTCTCTCAAGGGTCTCATGGCAGGAGAACCGCACGCCAACATCTGTCA	3287
Qy	3138	TCATCTGGTCCATCTCTGCTGGCTCCATCTTCTCCCTGCTGTGGGTGAGATCGACCTT	3197
Db	3288	TCGTTTGGTCCATCCTCTCTTGGGTATCTTCTCTCTGCTGGGTGAAGATCGATCCTT	3347
Qy	3198	TCATCTGTGAGAC---CAAGGGCCCCGACGTCAGGCAGTGTTGGCATCAATTGCTGAGC	3252
Db	3348	TCATCTCCCGACACAGAAAGCTCCGCCCTTGGGGCAATGTGTGTCAATGTCTGATC	3405

RESULT 11

```

RESULTS 11
US-10-160-719A-21
; Sequence 21, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthase
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 3725
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (179) ... (3400)
US-10-160-719A-21

```

[illegible]

Db 1614 CTGATGGAAGTCTGGCTGGGAATAAACCTAGGACCAATCTCGGCATGATTCAGGTGT 1673
Qy 1572 TCTTGGGCCACAGCGCGGCCACACAGACACCGAGGCAACGAGCTGCCCCGCTCGTGTACG 1631
Db 1674 TCTTGGGGCACAGTGGTGGGCTTGACACTGATGGAATGAATTAACACAGCTCTTGTCTATG 1733
Qy 1632 TCTCCCGTGAAGCGCCCGGGAATTCAGACACACAGAAAGCGCGGCATGAACGCTC 1591
Db 1734 TCTCTCGTGAAGAGACCAAGGCTTCAGCATCACAGAAAGCGTGGTGAATGAATGCAC 1793
Qy 1692 TGATTCGCGTCTCCGCGCTGCTGACCAACGCGCATTCATGCTCAACTTGGACTGTGATC 1751
Db 1794 TGATTCGCTGATCTGCTGCTGACAAATGGTGCCTATCTTCAATGTGGAATGTGACC 1853
Qy 1752 ACTACATCAACAAACAGCAAGCCCATCCGGGAGGCGCATGTCTCTCATGGAACCTCAGG 1811
Db 1854 ATTACTTCAATAGCAGCAAGCTCTTAGAAGCAATGTCTTCATGATGATCCAGCTC 1913
Qy 1812 TCGCGCGGAGGTCTGCTAGCTTCAGTTCAGTTCGCGAGAGGTTGACGCGCATCGACGTGACG 1871
Db 1914 TAGGAAGGAATACTTGTATGTATCAATTTCCACAAAGATTTGATGGCAATTGACTTGCAAG 1973
Qy 1872 ACCGATACGCTAAACAGGAACACCGTCTTCTTCGACATCAACATGAAGGGGCTCGACGGCA 1931
Db 1974 ATCGATATGCTAATAGGAACATAGTCTTCTTGATATCAACATGAAGGTCTAGATGGCA 2033
Qy 1932 TCCAAGCCCGGTGTAGCTCGGACAGGGTGCCTGTTCCGGCGCCAGGCGCTCTACGGCT 1991
Db 2034 TCCAGGCTCCAGTCTATGTGGGAACAGGATGCTGTTTCAATAGCAGGCTTTGTATGGAT 2093
Qy 1992 ACAACCT-----CCCAAGGACCCNAGAGGCCCAAGATGGTGACCTGGCATGTGCC 2045
Db 2094 ATGATCCTGTTTGAAGTCTGATCTGGAACCTTAACATTTGTTAAGAGCTGCTGTG 2153
Qy 2046 CGTGTCTCGGCCGCAAGACGGA----- 2069
Db 2154 GTAGAGGAGAGAGAGAGACAGATTAATGATAGTCAAGCGCTATTATGAAGAA 2213
Qy 2070 -----AACAGCCAAAGGACGGGCTCGGAGGACCGCTGATATGGGAGTAG 2117
Db 2214 CAGAACTTTCAGCTCCCATCTTTAAACATGAAGACATCGAGGAGGTATTGAAGGTTATG 2273
Qy 2118 ATAGCAGCAGGAGATGCTCATGTCCACATGAATTCGAGAGCGGTTCCGCGAGTCCG 2177
Db 2274 AGGATGAAGGTCAGTGTCTTATGTCCAGAGGAAATTGGAGAAACGCTTTGGTCACTCTC 2333
Qy 2178 CGCGGTTCTGCTCAGCTGACGCTGATGAGGAAGCGCGCTCCCTCTCGTCGAGCCCG 2237
Db 2334 CAATCTTCATGATCCACCTTTATGACTCAAGTGGCATACCACCTTCAACAAACCCAG 2393
Qy 2238 CGCGCTCTCAAGGAGGCCATCCATGTCTACAGCTCGGCTACGAGGACAAACCGACT 2297
Db 2394 CTCTCTACTGAAGGAGCTATCCATGTTATCAGCTGTGGGTACGAGGACAAACTGAAT 2453
Qy 2298 GGGGGCTGGAGCTGGGTGATCTACGGTGCATTCAGGAGGACATCTGACCGGGTTCA 2357
Db 2454 GGGGAAAGAGATTGGCTGATCTATGGTTCAAGTACAGAGGATTTCTGACTGGGTTA 2513
Qy 2358 AGATGCACTCCCGGGTGGCGCTCGGTGACTGATGCGGAAGCGGGCGGTCAAGG 2417
Db 2514 AATATGATGCAAGAGGTGCAATCAATCTACTGATGCCACACGACCTTTGTTCAAGG 2573
Qy 2418 GGTCCGCGCGCATCAATCTATCGGACCGTCTCAACAGAGTGTCTCCGGTGGCGCTGGGT 2477
Db 2574 GTTCGACCAATCAATCTTCTGATCGTCTTAATCAGGTGCTCCGTTGGGCTCTTGGGT 2633
Qy 2478 CGCTCGAGATCTTCTCAGCGGACAGCCCCCTGTGTAGGCTACAGAAACCGCAACC 2537
Db 2634 CAGTGGAAATCTGCTTAGCAGACATTTGCTCTATATGTTATGCTACAAAT---GGGCGAT 2690
Qy 2538 TCAAGTGGCTGGAGCGCTTCGCTTACATCAACACCAACATCTACCCCTTCACCTCGCTCC 2597
Db 2691 TGAAGCTTTGGAGAGGCTGGCTTACATTAACACCATTTGTTATCCAAATCAGATCTGTTT 2750

Qy 2598 CGTGTCTCGCTACTGCAACCTCCCGCGTCTGCTCTCTCAACCGGCAAGTTTCATCATGC 2657
Db 2751 CGCTTATCGCTATTGTGTCTCTCTGCTATCTGTCTTACCAATAAATTTATCATTC 2810
Qy 2658 CGTCGATTAAGCAGTTTCGCCAGGCTCTTCTTCATCGCCCTCTTCATGTCCATCTTCGCGA 2717
Db 2811 CTGAGATTAATATGCTGGAAATGTTCTTCAATCTTCTTTTGGCTCCATTTTCGCAA 2870
Qy 2718 CGGGCATCTCGAGATGCGGTGGAGCGGTGAGCATCAGAGAGTGGTGAGGAACGAGC 2777
Db 2871 CTGCTATATTGAGCTCAGATGAGTGGTGGCAATTGAAGATTGGTGAGGAATGAGC 2930
Qy 2778 AGTTCCTGGTCTATCGCGGCGGTTCGCGCATCTCTTTCGCGCTGTCGAGGGCTTGCTCA 2837
Db 2931 AGTTTTGGGTATTGGTGGCACTCTCGCCCACTCTTCGCGGTGTTCCAGGGTCTGCTGA 2990
Qy 2838 AGTTCCTCGCGGGATCGACACCACTTCACCGTCACTCCCAAGGCCACCGGCGACGAGG 2897
Db 2991 AAGTGTGGCTGGATTGATACCAACTTCACAGTTACCTCAAGGCATC---TGATGAGG 3047
Qy 2898 ACAGCAGTTCGCGGAGCTCTAGCCCTTCAAGTGGACACGCTCTCTCATCCCGCCACCA 2957
Db 3048 ATGGCGACTTTGCTGAGCTATATGTTCAAGTGGACCAAGTTGCTCATCCCTCCGACCA 3107
Qy 2958 CGTGTCTCATATTAACGTCTATCGCGCTCGTGGCCCGGCATCTCCGACGCCATCAACAAAG 3017
Db 3108 CTGTTCTGTCTATTAACCTGGTCCGAATGGTGGCAGGAATTCGTATGCCATTAAACAGCG 3167
Qy 3018 GGTACCAAGTCTGGGGGCCCCCTCTTCGCGCAAGCTCTTCTCGCTTCGGGTTCATCGTCC 3077
Db 3168 GCTACCAATCTGGGCTCCGCTCTTTGGAAGCTGTTCTTCGATCTGGGTGATCCTCC 3227
Qy 3078 ACCTCTACCGTTCCTCAAGGGCTCATGGGGCGCCAGAACAGAGCGCCACCGTTGTTG 3137
Db 3228 ATCTCTACCCCTCTCTCAAGGCTCTCATGGGAGGAGGAGGAGCGCCACCAATCGTCA 3287
Qy 3138 TCATCTGTGCTCATTTCTGTGGCTCCATCTTCTCTCTGCTGTGGGTGAGGATCGACCTT 3197
Db 3288 TCGTTTGGTCCATCTCTCTTGGCTCTATCTTCTCTGCTGTGGGTGAAAGATCGATCCTT 3347
Qy 3198 TCATCTGCTAGGAC---CAAGGGCCCGGAGCTCAGGCAAGTGTGCGATCAATTTGCTGAGC 3252
Db 3348 TCATCTCCCGACACAGAAAGCTGCGCCCTTGGGCGCAATGTGTGTGAACCTGCTGATC 3405

RESULT 12
US-10-160-719A-41
; Sequence 41, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 3725
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (179)...(3400)

US-10-160-719A-41

Query Match 34.5%; Score 1187.2; DB 4; Length 3725;
Best Local Similarity 62.8%; Pred. No. 1.3e-210;
Matches 2033; Conservative 0; Mismatches 1123; Indels 82; Gaps 9;

Qy	89	GCGATGAGGCGCAGCGCGGCTGTCGCGGCTCGCACACCGGACGAGCTGGTGTG	148
Db	176	GAGATGGCGGCAACAGAGGATGGTGCAGGCTCTCACACCGCAGAGTTCGTGATG	235
Qy	149	ATCCGGGCGCACGAGGACCC- - - - - AAGCCGCTGCGGGGCTGAGCGGCGAG	196
Db	236	ATCCGCCACGACGCGCGCCTCCGCGCTAAGCCACGAGAGTGCGAATGGCGAG	295
Qy	197	GTGTGCGAGATATCGGCGACGAGTGGGCTCAAGTGGACGCGACCTCTTCGTGCGC	256
Db	296	GTCTGCCAGATTTGTGGCGACACTGTTGGCGTTTCAGCCACTGGTGATGCTTTGTTGCC	355
Qy	257	TGCAACAGTGGGCTTCCCGTGTGCGGCGCTTACGAGTACGAGCGCGGAGGCG	316
Db	356	TGCAATGAGTGGCTTCCCTGTCTGCGCGCTTGTATGATGATGAGTACGAGCGCAAGG	415
Qy	317	ACGCAGAACTGCCCCCGAGTCAAGACGCGCTTCAAGCGCTCAAGGGAGCGCGAGGTT	376
Db	416	AACCAATGCTGCTCAGTGCAGACTAGATACAGAGACAGAAAGTAGCCCTCGAGTT	475
Qy	377	GCGGGGACGATGACGAGGAGCAATCGACGCTGGAGCACGAGTTCAACATCGACAC	436
Db	476	CATGCTGATGATGAGGAGGAAGATGTTGATGACCTGGACAAATGAATCAACTAAGCAA	535
Qy	437	GAGAAATCAGACGAGCTGGAGGCGCAATGACGACAGACCCAGATCACGAGCGCATG	496
Db	536	GGCAATGGGAAGGGCC- - - - - CAGAGTGGCAGCTTCAAGGAGATGACGCTGATCTTTCAT	593
Qy	497	CTGCACGCGAGATGAGCTACGAGAGGGCGCCGACGACGCGCGCAACACACCCCG	556
Db	594	CTGCTGCCATGACCCACACCATCGGATTCCAGCCCTTACAGTGGACACAGATATCTG	653
Qy	557	CAGATCCCGC- - - - - CATCATCACCGGCTCCCGTCCGTCGCGTGGAGCGGTGATTTCC	612
Db	654	GAGAGATCCCTGATGATCCCTGACCGTCACTTCTATCCGCGATCCAAATCGAGCTATG	713
Qy	613	GATTAACCAACGGGTATGGCACGCGAGGTCTGCTTCTTCCGTGCACAGCGCATCCATCC	672
Db	714	TTGATCAACGGTTCAGATTCCTGTGAGGATTTGGACCCCTCGAAGGACTTGAATTCCT	773
Qy	673	GTACCTGTGTCTGAGCGAGGAGTGCCAGTGGGAGAGAAAGAGTGGAGCTG- GA	731
Db	774	ATGGGCTTAATAGTGTGACTGGAAGGAAGAGTGGAGCTGAGGGTTAAACAGGACA	833
Qy	732	AGGAGGATGAGACGACTGGAAGTCCAAAGCGGCGATCCTCGCGCGGCGCGGATCCCG	791
Db	834	AAAATATGTTGCAAGTGACTAATAAATATCCAGAGGCTAGAGGAGACATGGAGGGACTG	893
Qy	792	AAGACATGGAGCGCGAGTGCGACTGAACGACGAGGCGAGCGACCGCTGTGAGGAGG	851
Db	894	GCTCAAAATGGAGAGATATGCAAAATGGTTGATGATGACGCGCTACCTTTGAGCGCGATTG	953
Qy	852	TGTCGATCGCTCGAGCAAGGTGAACCGTACCGGATGGTATCGTGGTGGCTCGCTTG	911
Db	954	TGCCAATTTCTCAACAGCTCAACCTTTACCGGATAGTAAATCATTTCCGCTTTATCA	1013
Qy	912	TGCTCGGCTTCTTCCCTCGGTACCGTATCTTCGACCCCGCTCCCGAGCGCCATCGGGCTGT	971
Db	1014	TCCTGTGCTTCTTCTTCAATATCGTATCAGTATCATCCAGTGGTAAATGCTTATGATGT	1073
Qy	972	GGCTCGTCTCATCATCTGCGAGATCTGGTTGGCCATCTCTGTGATCTCTGACCAAGTTCC	1031
Db	1074	GGCTAGTATCTGTTATCTGTGAGGTCTGGTTTGGCTTGTCTGCTGCTTAGATCAGTTCC	1133
Qy	1032	CCAAGTGTTCCTCCATCGACCGCAGAGTACCTCGACCGGCTCTCCCTCAGGTACGAGA	1091
Db	1134	CAAAATGGTATCCAATCAACCGTGAGACATATCTCGACAGGCTTGCATTTGAGGTATGATA	1193

Qy	1092	GGGAAGGGAGCCGCTGCTGCTGTCGGGGTGGACCTGTTCTGAGCACCGTGGACCCGC	1151
Db	1194	GAGAGGAGAGCCATCAAGCTGGCTCCCATTTGATGTTGTCAGTACAGTGGATCCAT	1253
Qy	1152	TCAGGAGCGCGCTGTCAGCCCAACACCGTGTCTCCATCTCGCCGTAGACTACC	1211
Db	1254	TGAAGGAACCTCAGTATCAGCCCAACTGTTTGTCTCATCTTCTGCTGTGATACC	1313
Qy	1212	CCGTGGCAAGGTTCTCTGCTTACGTCGACGCGCGGTGATGCTGATCGTTCGAGT	1271
Db	1314	CTGTTGACAAAGTGTCTATGTTTCTGATGATGCTCAGCTATGCTGATTTTGTAGT	1373
Qy	1272	CGCTGTCGAGACGCGCGGAGTTCCGCGCGAAGTGGTGGCTTCTGCAAGAGTTCGGCA	1331
Db	1374	CTCTCTCTGAACTGCGCGAATTTGCTAGAAAGTGGTTCCTTTTGAAGAGCAATA	1433
Qy	1332	TCGAGCCCGCGCGCGGAGTTCTACCTTCTCGCTCAAGGTGCTACTCTCAAGGACAAAG	1391
Db	1434	TTGAACCAAGAGCTCCAGAAATTTTACTTTGCTCAAAATAAGATTTACTGAGGACAAA	1493
Qy	1392	TGCAGCCACCTTCTGTCAGGAGCGCGCCATGAAGAGAGATATGAGGAGTTCAAGG	1451
Db	1494	TTCAACCTTCAATTTGTTAAGGAAGACGAGCAATGAAGAGAGATATGAAGAAATTCAAA	1553
Qy	1452	TCGGAATCAACGCGCTGGTGGCCCAAGGCCATGAAGGTGCCGCGAGAGGGGTGATGATGA	1511
Db	1554	TAAGAAATCAATGCTTGTTCGCAAGACAGAAAGTGCCTGAAAGAGGGGTGACCATGG	1613
Qy	1512	AGGAGGACGCGCTGGCGGGAACAAACCCGCGACCAACCCGCGCATATCCAGGTGT	1571
Db	1614	CTGATGGAATCTGCTTGGCTGGGATTAACCTTAGGACCATCTCTGGGATGATTCAGGTGT	1673
Qy	1572	TCCTGGGCCACAGCGCGGCCACGACACCGAGGGCAACGAGCTGCCCGCTCGTGTACG	1631
Db	1674	TCCTGGGGCACAGTGGTGGCTTGACACTGATGGAATGAATTAACACGCTCTTGTCTATG	1733
Qy	1632	TCCTCCGTGAGAGCGCGCGGATTCAGACACCAAGAGAGCGCGCGCCATGAACGCTC	1691
Db	1734	TCCTCTGTAAGAGAGACGAGGCTTTCAGCATCAAGAAAGGCTGCTGCAATGAATGCAC	1793
Qy	1692	TGATTCGCTCTCGCGCTGCTGACCAACCGCCATTCATGCTCAACTTGGACTGTGATC	1751
Db	1794	TGATTCGTGATCTGCTGTGCTGACAAATGCTGCTATCTTCTCAATGTGATTTGAC	1853
Qy	1752	ACTACATCAACAAACAGCAAGCCATCCGAGGAGCCATGTGCTTCTCATGAGACCTCAGG	1811
Db	1854	ATTACTTCAATAGCAGCAAGCTCTTAGAGAGCAATGTGCTTCAATGATGATCCAGCTC	1913
Qy	1812	TCGCGCGAAGGTTCTGCTACGTTTCAGTTCGCGCAGAGGTTGACGCGCATTCGATGACG	1871
Db	1914	TAGGAAGAAAACCTTGTATGTACAAATTTCCAAAGATTTGATGGCATTCGACTTGCACG	1973
Qy	1872	ACGATACGCTACAGGAACACCGCTCTTCTCGACATCAACATGAAGGGGTGACGCGCA	1931
Db	1974	ATCGATATGCTAATAGGAACATAGTCTTCTTGTATATCAACATGAAGGCTTAGATGGCA	2033
Qy	1932	TCCAAGCCCGGTGTACGTCGGACAGGCTGCGTGTTCGCGCGCAGCGCTCTACGCGCT	1991
Db	2034	TTGAGGTCAGTCTATGTGGAAACAGATGCTGTTTCAATAGGACGCTTTGTATGAT	2093
Qy	1992	ACAAACCT- - - - - CCAAGGAGCCCAAGAGGCCAAGATGGTGAACCTGCGACTGTGCC	2045
Db	2094	ATGATCCTGTTTGTGACTGAAAGTGTCTGGAACCTAACAATGTTGTTTGAAGAGCTGCTGTG	2153
Qy	2046	CGTCTTCGGCGCCCAAGAGCGG- - - - -	2069
Db	2154	GTAGAAGAAAGAAAGAAACAAGATTTATGGAATTCAAAGCCGCTATTTATGAAGAA	2213
Qy	2070	- - - - - AACACGCCAAGAGCGGCTGCGGAGGGCACCGCTGATATGGGAGTAG	2117
Db	2214	CAGAACTTTCAGCTCCCATCTTTAATGGAAGACATCGAGGAGGTTATTCAGGTTATG	2273

Qy	613	GATTACCAACGGGTATGGCCACGGGAGGTCTCGTCTTCTCCCTGCACAAAGCGCATCCATCC	672
Db	719	TTGATCAAAAGCGTCCCAGTTCTCTGTGAGGATTGTGGACCCCTTCCAAAGGACTTGAATTCCT	778
Qy	673	GTACCCCTGTGTCGAGCCAGGAGTGCCCAAGTGGGACGAGAAGAAAG	728
Db	779	ATGGGCTTAAATAGTGTGACTGGAAAGAAAGAGTTGAGAGCTGGAGGGTTAAACAGGACA	838
Qy	729	GGAAAGAGAGGATGGACGACTGGAAAGTCCAAAGCAGGGCATCTCGGGGGGGCCGCCGATC	788
Db	839	AAAATATGATGCAAGTGAATTAATATCCAGAGGCTAGAGGAGGAGACATGGAGGGGA	898
Qy	789	CCGAAGACATGGACCGCGACGTGGCACAATAAGCAGAGAGGCGAGCAGCCGCTGTGAGGA	848
Db	899	CTGGCTCAAAATGGAGAGATGACAAATGGTGTGATGATGCACGGCTACCTTTGAGCCGTA	958
Qy	849	AGGTGTCGATCGCGCTCGAGCAAGGTGAACCGGTACCGGATGGTGCATCGTGGTGGCTCTCG	908
Db	959	TCGTGCCAAATTTCTCAAAACAGCTCAACCTTTACCGGGTAGTAGTCAATCTCCGCTCTTA	1018
Qy	909	TTGTGCTCGCCTTCTTCTCCGTACCGTATCTCTGCACCCCGTCCCGGAGGCCATCCGGC	968
Db	1019	TCATCTGTGCTTCTTCTTCAGTATCGTGTCACTCATCCAGTGGGTGATGTATGAT	1078
Qy	969	TGTGGCTCGTCTCATCATCTGCAGATATGTGTTGCGCATCTCTGTGATCTCTGACCACT	1028
Db	1079	TATGGCTAGTATCTGTATCTGCGAGGCTCTGGTTTGGCTTGTCTTGGCTTCTAGATCAGT	1138
Qy	1029	TCCCCAAGTGGTTCCCATCGACCGCAGACGTACTCTGCACCGGCTCTCCCTCAGGTACG	1088
Db	1139	TCCCAAAATGGTATCCAAATCAACCGTGAGACATATCTTGAACGGCTTGCATTAGGTATG	1198
Qy	1089	AGAGGGAAGGGAGCCGCTCGCTGTCTGGCGGTGGACCTGTTCTCGTGAGCACCGTGGACC	1148
Db	1199	ATAGAGAGGAGAGCCATCACAGCTGGCTCCCATTTGATGTCTTCTGTCAGTACAGTGGATC	1258
Qy	1149	CGCTCAAGAGACCGCGCTGGTGACGCCCAACACCGTGTCTCTCATCTCGCCGTAGACT	1208
Db	1259	CAITGGAAGGAACCTCCACTGATCACAGCCAACTGTTTGTCCATTTCTTCTGTGGATT	1318
Qy	1209	ACCCGTGGACNAGTCTCTCTGCTACGCTCTCCGACGAGGGGGCTCATCTGACGCTTCG	1268
Db	1319	ACCTGTTTGAAGAAGTGTATGCTATGTTTCTGATGATGGTTTCACTGCTATGCTGACTTTG	1378
Qy	1269	AGTCGCTGTCTGGAGACGGCCGAGTTCTCGCGGCAAGTGGGTGCCCTTCTGCAAGAAGTTCG	1328
Db	1379	AGTCTCTCTCAGAAACCGCAGATTGCTAGAAAGTGGTTCCCTTTGTAGAAGACACA	1438
Qy	1329	GCATCGAGCCCGCGCCCGGAGTTCTATTCTCGCTCAAGGTGCACTACCTCAAGGACA	1388
Db	1439	ATATTGAACCAAGAGCTCCAGAAATTTTACTTTGTCTCAAAAATAGATTACTGAAAGGACA	1498
Qy	1389	AGGTGACGCCACCTTCGTGCGAGAGCGCGGCCCATGAGAGNAGAGTATGAGGAGTTCA	1448
Db	1499	AAATTCAACCTTCAATTTGTTAAGGAAGAAGCGCAATGAAGAGGGAGTATGAAGAAATCA	1558
Qy	1449	AGGTCCGGATCAACCGCTGTGGCCAAAGGCCATGAAGGTCCGGCAGAGGGGTGGATCA	1508
Db	1559	AAGTAAAGATCAATGCCCTTGTGGCCAAAGCAGAAAGTGCCTGAAGAGGGGTGACCA	1618
Qy	1509	TGAAGGACGGAACCGCTGGCCGGGAAACAACCCCGCAACCAACCCCGGATGATCCAGG	1568
Db	1619	TGGCTGATGAACTGCATGGCTGGGAAATAATCTAGGGACCATCTCTGGCATGATTCAGG	1678
Qy	1569	TGTTCTTGGGCGCACAGCGGGCGGCCACACACCGAGGGCAACGAGCTGCCCGGCTGTGT	1628
Db	1679	TTTTTCTTGGGGCACAGTGGTGGGCTGCACACTGATGGAAATGAGTTTACCAACGCTTGTCT	1738
Qy	1629	ACGTCTCCGTGGAAGCGCCGGATTTCTCAGACACCAAGAGAGGGCGCGCCATGAACG	1688
Db	1739	ATGTCTCTCGTGAAGAGACCAAGGCTTTCAGATCAAGAGAGGCTGGTGCATGAATG	1798
Qy	1689	CTCTGATTCGCTCTCCGCGGTGTGACCAACCGCCATTCATGCTCAACTTGGACTGTG	1748

[illegible]

```
Db 2876 CCACTGGTATATTGGAGCTTAGATGGAGTGGTGTGGCATTGAAGATTGGTGAGAAATG 2935
QY 2775 AGCAGTTCTGGGTATCGCGGGGTGTCCGCGCATCTCTTCGCGCGTGTGCAGGGGCTGC 2834
Db 2936 AGCAGTTTGGGTATTGGTGGCACCTCTGCCCCATCTCTTCGAGTGTTCACAGGTCTGC 2995
QY 2835 TCAAGTCTCGCGGGGATCGACCAACATTCACCGTACCTCAAGGCCACCGGGGACG 2894
Db 2996 TGAAGTGTGGCTGGGATTGATACCAACTTCAAGTTACCTCAAAGGCATC---TGATG 3052
QY 2895 AGGACGACGATTGCGCGAGCTCTACGCTTCAAGTGGACCACTCCTCATCCGCCCA 2954
Db 3053 AGGATGGGACTTTGCTGAGCTATATGTGTTCAAGTGGACCAAGTTGCTCATTTCTCCGA 3112
QY 2955 CCACGCTGCTCATATTAACTGATCGGCGTCTGGCGGCGCATCTCCGACGCCATCAACA 3014
Db 3113 CCATGTTCTTGTCAATTAACCTGGTGGATGTGGCAGAAATTTCTATGCCATTAAACA 3172
QY 3015 ACGGGTACCAAGTCTCGGGGCCCTCTTTCGGCAAGCTCTTCTTCGCTTCGGGTCAATCG 3074
Db 3173 GTGGCTACCAATCTCGGGGTCCGCTCTTTGGAAAGCTGTTCTTCGATCTGGGTGATCC 3232
QY 3075 TCCACCTCTACCGTTCCTCAAGGGGCTCATGGGGCCGAGACAGAGCCGCCACCGTTG 3134
Db 3233 TCCATCTCTACCCCTTCTCAAGGGTCTCATGGAAAGGCGAAGCCGACACCAACATCG 3292
QY 3135 TTGTCATCTGGTCCATTTCTCGGCTTCCATCTCTTCTCCCTGCTCTGGGTGAGATCGACC 3194
Db 3293 TCATTGCTGTCATCTCTTTCGATCATCTCTCTCTGCTGTGGGTGAAGATCGATC 3352
QY 3195 CTTTCATGCTCAGAC---CAAGGGCCGGAOCTCAGGCAAGTGTGGCATCAATGCTGAG 3251
Db 3353 CTTTCATCTCCCGACACAGAAAGCTGCTGCTTGGGGCAATGTGGGTCAACTGCTGAT 3412
QY 3252 C 3252
Db 3413 C 3413

RESULT 14
US-10-160-719A-53
; Sequence 53, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE OF INVENTION: Thereof
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 3753
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (184)...(3408)
US-10-160-719A-53

Query Match 34.3%; Score 1180.2; DB 4; Length 3753;
Best Local Similarity 63.0%; Pred. No. 2.6e-209;
Matches 2043; Conservative 0; Mismatches 1113; Indels 85; Gaps 11;

QY 89 GCGATGAGCCAGCGCGGGTGGTGGCGGCTCGCACAAACCGGAACGAGCTGGTGCTG 148
```

```
Db 181 GAGATGCGCGCCCAACAAAGGGGATGGTGGCGGCTCGCACAAACCGAGTTTCGTCAATG 240
QY 149 ATCCGGGGCCACGAGGACC-----CCAAGCCGCTGCGGGCGCTGAGCGGCGAG 196
Db 241 ATCCGGCCACGACGGCGATGTGCGGGCTCGGCTTAAGCCCAAAAGAGTGGGAATGACAG 300
QY 197 GTGTGCGAGATATGCGGCGACGAGGTGCGGCTCACGGTGGACGCGGACCTCTTCGTGCGC 256
Db 301 GTCTGCCAGATTTCGCGTGACTCTGTGGGTGTTTACGCCACTGGTGATGTCTTTGTGCC 360
QY 257 TGAACGAGTGGCGCTTCCCGTGTCCCGCTCTGTACGAGTACGAGGCCCGGGAGGC 316
Db 361 TGAATAGTGTGCTTCCCTGTCTGCGGCCATGTATGAGTATGAGCGCAAGGAGGG 420
QY 317 ACGCAGAAATGCCCCCAGTGCAGACGCGCTCAAGGCGCTCAAGGGGAGCCCGAGGTT 376
Db 421 AACCAATGTGCCCCCAGTGCAGACTAGATACAAGAGACAGAAAGGTAGCCCTCGAGTT 480
QY 377 GCCGGGACGATGACGAGGAGGACATCGACGACTCGAGCACGAGTTCAACATCGACGAC 436
Db 481 CATGGTGTAGGATGAGGAGATGTTGATGACTTAGACAAATGAATTTCAACTACAGCAA 540
QY 437 GAGAATCAGAGGAGGCTGGAGGGCAACATGCAGAACAGCCAGATCACCGAGGCGATG 496
Db 541 GGCAGTGGGAAAAGGCC--CAGAGTGGCAACTGCAAGGAGATGATGCTGATCTCTCTTCAT 598
QY 497 CTGCAAGGAGGATGAGCTACGGGAGGGGCGCGACGCGGACGCGGCAACACACACCCCG 556
Db 599 CTGCTCGCCATGAGCCACATCATCGGATTTCCAGCCTCGAACGCGTCAACAGATATCTG 658
QY 557 CAGATCCCCCCTCATCATCACCGGCTCCCGCTCCGT-----GCCGGTGAAGCGGTGAGTTTCC 612
Db 659 GAGAGATTCTGATGCTTCCCTGACCGTCAATTCATCCGAGTCCCAACATCGAGTATG 718
QY 613 GATTACCAACGGGTATGGCCACGCGAGAGTCTGCTTCTCCCTGCAACAGGCGATTCATCC 672
Db 719 TTGATCCAAAGCGTCCCAAGTTCTCTGTGAGGATTTGAGACCCCTCGAAGGACTTGAATTCCT 778
QY 673 GTACCTGTGTCTGACGAGGAGTSCCAAGTGGGACGAGAGAAAG---AAGTGAGCT 728
Db 779 ATGGGCTTAATAGTGTGACTGGAAGAAAGATTGAGAGCTGGAGGGTTAAACAGACA 838
QY 729 GGAAGGAGGAGTGGACGACTGGAAGTCCAAGCAGGGCATCTCTCGGCGGCGCGCATC 788
Db 839 AAAATATGATGCAAGTGAATAATAATATCCAGAGCTAGAGGAGGAGACATCGAGGGGA 898
QY 789 CGAAGACATGAGACGCGGCTGGGCTGAAACGACGAGGAGGAGCGGCTGTCTGAGGA 848
Db 899 CTGGCTCAAATGGAGAAGATATGCAAAATGGTTGATGATGACGCGCTACCTTTGAGCCGTA 958
QY 849 AGGTGTGATCGGTCGAGCAGGAGTGAACCGGTACCGGATGGTGTGATCGGTGCGTCTCG 908
Db 959 TCGTGCCAAATTTCTCAAACCGAGCTCAACCTTTACCGGGTGTGATGATCTCTCCGCTTA 1018
QY 909 TTGTGCTCGCTCTTCTCTCCGTAACCGTATCTGACACCCCGTCCCGGAGCGCATCGGGC 968
Db 1019 TCATCTGTGCTTCTTCTCCAGTATCGTGTGATCATCCAGTGTGCTGATGCTTATGAT 1078
QY 969 TGTGGTCTGTCTCCATCATCTGCGAGATCTGGTTCGCCATCTCTCTGATCTCTCGACGAT 1028
Db 1079 TATGGCTAGTATCTGTTATCTGCGAGGTCTGGTTTGCCTTGTCTTGGCTTCTAGATCAGT 1138
QY 1029 TCCCCAAGTGGTTCCCATCGACGAGAGCTGCTGAGCGGCTTCCGCGGCTCTCCCTCAGGTACG 1088
Db 1139 TCCCAAAATGGTATCCAATCAACCGGAGACATATCTTGACAGGCTTGCATTTGAGGTATG 1198
QY 1089 AGAGGGAAGGGAGCGCTGCTGCTCGCGGCTGGACCTGTTCTGAGCAGCACCGTGGACC 1148
Db 1199 ATAGAGAGGAGAGCCATCACAGCTGGTCCCAATTGATGTCTTCGTGCTGAGTACAGTGCATC 1258
QY 1149 CGTCTAAGAGCGCGGCTGGTGAACGCCCAACCGGTGCTCTCCATCTCTCGCGGTAGACT 1208
```



```
RESULT 15
US-10-160-719A-9
; Sequence 9, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 3780
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)...(3425)
US-10-160-719A-9

Query Match      34.3%; Score 1180.2; DB 4; Length 3780;
Best Local Similarity 63.0%; Pred. No. 2.6e-209;
Matches 2043; Conservative 0; Mismatches 1113; Indels 85; Gaps 11;

QY      89  GCGATGAGGCGAGCGCGGCTGTGGCGGCTCGCACACCGGACGAGCTGTGCTG 148
DB      198  GAGATGGCGGCCAACAAAGGGGATGGTGGCGGCTCGCACACCGCAACGAGTTCGTCATG 257

QY      149  ATCCGGGGCCACAGGAGACC-----CCAAAGCCGCTGGCGGCGCTGAGCGGGCAG 196
DB      258  ATCGGCCACACGGCGATGTGCCGGGCTCGGCTAAGCCCAACAAAGATGCGAATGACAG 317

QY      197  GTGTGGAGATATGCGCGGACGAGGTGCGGCTCACGGTGAACCGCGACCTCTTCGTGCGC 256
DB      318  GTCTGCCAGATTTCGCGTGACTGTGGGTGTTCAGCCACTGGTATGATCTTTGTGCC 377

QY      257  TGCACAGAGTGCCTTCCCGGTGTCGCGCCCTGTCTACAGTACGAGCGCGGGAGGGC 316
DB      378  TGCATAGTGTGCTTTCCCTGTCTGCCGCCCATGCTATGAGTATGAGCGCAAGGAGGGG 437

QY      317  ACGCAGAACTGCCCCCAAGACGCGCTACAAGCGCTCAAGGGGAGCCCGAGGGTT 376
DB      438  AACCAATGCTGCCCCAGTGCAGACTAGATACAAGAGACAGAAAGTAGCCCTCGAGTT 497

QY      377  GCCGGGACATGACGAGGAGGACATCGACGACTGGAGCACGAGTTCAACATCGACGAC 436
DB      498  CATGGTGATGAGGATGAGGAAGATGTTGATGACCTAGACAATGAATCAACTACAAGCA 557

QY      437  GAGAAACAGCAGAGGAGCTGGAGGCAACATGCAGACAGCCAGATCAACGAGGCGATG 496
DB      558  GGCAGTGGGAAAGGCC--CAGATGGCAACTGCAAGGAGATGATGCTGTCTTTCAT 615

QY      497  CTGCAACGCGAGGATGAGCTACGGGAGGGGCCCCGACGACGCGGACGCGCAACAAACCCCG 556
DB      616  CTGCTGCCATGAGCCACATCATCGATTCCACGCTGACAGCGGTCAACAGATATCTG 675

QY      557  CAGATCCCGCCATCATCAACGGGTCCCGTCCGT-----GCCGGTGAAGCGGTGAGTTTC 612
DB      676  GAGAGATTCTGATGCTTCCCTGACCGTCAATCTATCCGAGTCCCAACATCGAGCTATG 735

QY      613  GATTACCAACGGGTATGGCCACGCGGAGTCTCGTCTTCCCTGCACAAAGCGCATCCATCC 672
DB      736  TTGATCCAGCGTCCAGTTCTCTGTGAGGATGTGGACCCCTCGAAGGACTTGAAATTCCT 795
```

```
QY      673  GTACCCCTGTGTCTGAGCCAGGAGTGCCAAAGTGGGACGAGAAAGAAAG-----AAGTGAGCT 728
DB      796  ATGGGCTTAATAGTGTGACTGGAAGGAAAGAGTTTGAGAGCTGGAGGGTTAAACAGGACA 855

QY      729  GGAAGGAGAGGATGGACGACTGGAAGTCCAAGCAGGGGATTCCTCGCGGGCGGGCCGATC 788
DB      856  AAAATATGATGCAAGTGAATAATAATATCCAGAGGCTAGAGGAGGAGAGACATGGAGGGGA 915

QY      789  CCGAAGACATGGACGCGGAGCTGGAACGACGAGGCGGAGGAGCGCTCTCGAGGA 848
DB      916  CTGGCTCAAATGGAGAAGATATGCAAAATGGTGTGATGATGACCGGCTACCTTTGAGCGGTA 975

QY      849  AGGTGTGATCGCGTCGAGCAAGGTGAACCGGTACCGGATGGTGTGATCGTGGTCTGCTCG 908
DB      976  TCGTGCCAAATTCCTCAAAACAGCTCAACCTTTACCGGGTAGTGATCATTTCTCGTCTTA 1035

QY      909  TTGTGCTCGCTCTTCTCCCGGTACCGTATCTCTGACCCCGTCCCGGAGCGCATCGGGC 968
DB      1036  TCATCTGTGCTTCTTCTCCAGTATCGTGTGATCATCCAGTCCGCTGATGCTTATGGAT 1095

QY      969  TGTGGCTCGTCTCCATCATCTCGAGATCTGGTTGCCCATCTCTGATCTCTCGACAGT 1028
DB      1096  TATGGCTAGTATCTGTATCTCGAGGTCTGGTTTGCCTTGTCTTGGCTTCTAGATCAGT 1155

QY      1029  TCCCCAAGTGGTTCCCATCGACCGGAGACGTACCTCGACCGCTCTCCCTCAGGTACG 1088
DB      1156  TCCCAAAATGGTATCCAAATCAACCGTGAGACATATCTTGACAGGCTTGCAATTCAGGTATG 1215

QY      1089  AGAGGGAAGGGAGCGCTGCTGTCTGGCGGTGACCTGTTCTGTGAGCACCGTGGACC 1148
DB      1216  ATAGAGAGGAGAGCCATCACAGCTGGCTCCCATTTGATGTCTTTCGTGAGTACAGTGGATC 1275

QY      1149  CGCTCAAGGAGCGCGCTGTGACCGCAACACCGTGTCTCCATCTCTCGCGGTAGACT 1208
DB      1276  CATTGAAAGAACCTCACTGATCAGACCAACACTGTTTGTCCATCTTCTCTGTGATTT 1335

QY      1209  ACCCGCTGGACAAAGTCTCTCTGCTCTCGACGACGCGCGCTCGATCTGACGCTTCG 1268
DB      1336  ACCCTGTGCACAAAGTGCATGCTATGTTTCTGTGATGATGGTTGAGCTATGCTGACTTTTG 1395

QY      1269  AGTCTGTGCGGAGCGCGGATTCGCGGCAAGTGGGTGGCTTCTGCAAGAGTTCG 1328
DB      1396  AGTCTCTCTCAGAAACCGCAGAAATTTGCTAGAAGTGGGTTCCTTTTGTAAAGAGCACA 1455

QY      1329  GCATCCAGCCCGCGCGCTCTACTTCTCGCTCAAGGTGCGACTACTCAAGGACA 1388
DB      1456  ATATTGAACAGAGCTCCAGAAATTTTACTTTTGTCTCAAAAATAGATTACCTGAGGACA 1515

QY      1389  AGGTGACGCCCCACTTGTGTCAGGAGCGCGGCCATGAAGAGAGAGTATGAGGAGTTCA 1448
DB      1516  AAATTCACCTTCATTTGTAAGGAAAGACGCGCAATGAAGAGGAGTATGAAGAATTCA 1575

QY      1449  AGGTCCGGATCAACGGCTGTGGCCCAAGCCATGAAGTTCGCGGAGAGGGGTGATCA 1508
DB      1576  AAGTAAAGATCAATGCTTGTTCGCAAGACACAGAAAGTGCCTGAAGAGGGGTGGACCA 1635

QY      1509  TGAAGGACGCGCAGCGCTGGCGGGAAACAAACCCCGCACACCCCGGCTGATCCAGG 1568
DB      1636  TGGCTGATGGAATGATGATGGCTTGGGAAATATCTTAGGACCATCTCTGGCATGATTCAGG 1695

QY      1569  TGTCTCTGGGCCACAGCGCGGCGCACGACACCGAGGGCAACGAGCTGCCCGCTCGTGT 1628
DB      1696  TTTTCTTGGGGCACAGTGGTGGCTCGACACTGATGGAAATGATTTACCACTGCTTGTCT 1755

QY      1629  AGTCTCCGCTGAGAAAGCGCGGGATTCAGACACACAGAAAGCGCGGCGCATGACG 1688
DB      1756  ATGTCTCTGTGAAAAGAGACCAAGGCTTTTCAGCATCAAGAAAGGCTGGTGCATTAATG 1815

QY      1689  CTCGTGATTCGCTCTCGCGGTGCTGACCAACCGCGCATTTTCATGCTCAACTTGGACTGTG 1748
DB      1816  CGCTGATTCGTGTATCTGTGTGTGACAAATGGTGCCTATCTTCTCAATGTGATGATGCG 1875
```

Qy	1749	ATCACTACATCAACAACAGCAAGGCCATATCGGGAGGCCATATGTGCTTCTCATGAGACCTTC	1808
Db	1876	ACCATTACTTTCAATAGCAGCAAAAGCTTTAGAGAAAGCAATATGTGCTTCATGATGATCCGG	1935
Qy	1809	AGGTCCGCCGGAGGTCTGCTAGCTTCAGTTCCGCGAGAGGTTCCGACGGCATCGACGTGC	1868
Db	1936	CTCTAGAAAGGAANAACTTGTATTGTACAAATTTCCACAGAGATTGTATGGCAATTTGACTTGC	1995
Qy	1869	ACGACCGATACGCTAAACAGGAACACCGTCTTCTTCGACATCAACATGAAGGGGCTGGACG	1928
Db	1996	ACGATCGATATGCTAATCGGAACATAGTTTTTCTTTGATATCAACATGAAGGCTCGGATG	2055
Qy	1929	GCATCCAAGGCCCGGTGTAGCTCGGGAACAGGTGCTGTGTTCCGCGCGCCAGGCGCTCTACG	1988
Db	2056	GCATTCAGGGTCCAGTTTACGTGGGAACAGGATGCTGTTTTCAATAGACAGGCTTTGTATG	2115
Qy	1989	GCTACAAACCTT-----CCCAAGGGACCCAAAGAGGCCCAAGATGTTGACCTGCGACTGCT	2042
Db	2116	GATACGATCTGTTTTCGACTGGAAGCTGATCTGGAGCCAAACATTTGTTATTAAAGAGCTGCT	2175
Qy	2043	G-----CCCGTGCCTTCGGCGCCGAAGAGCGGAAAAACACGCCAA	2079
Db	2176	GTGGTAGAAGGAAGAAAAGAAACAAGAGTTATTATGGATAGTCAAGCCGTATTATGAAGA	2235
Qy	2080	GGAGCGGCTGC-----CGAGGGGCAACGCTGATATGGAAGTAGATA	2120
Db	2236	GAACAGAAATCTTCAGCTCCCATCTTCAATATGGAAGACATCGAAGAGGGTATTGAAGGTT	2295
Qy	2121	GCACAGAGAGA-----TGCTCATGTCCACATGACTTCGAGAGAGCGGTTTCGGGCAGT	2174
Db	2296	ACGAGGATGAAGGTCAGTGCTTATGTCCAGAGGAAATTTGGAGAAACGCTTTGGTCAGT	2355
Qy	2175	CCGCGGCGTTTCGTCACTCGACGCTGATGAGGAAGCGCGCGTCCCTCTTCGTTCGAGCC	2234
Db	2356	CTCCTATTTCATTGCAATCCACCTTTATGACAACAAGTGGCATACCACTTCAACAACC	2415
Qy	2235	CCGCGCGCTCTCAAGAGGCCCATCATGTCACTACGTGCGGCTACGAGGAACAAGACCG	2294
Db	2416	CAGCTTCTCTACTAAAGGAAGCTATCCATGTCATCAGTTGTGATATGAGAACAAACATG	2475
Qy	2295	ACTGGGGCTGGAGCTGGGTGGATCTACGGGTGCGATCAGCGAGGACATCTCAGCGGGT	2354
Db	2476	AATGGGGAAGAAGATTGGCTGGATCTATGGTTCAAGTAACGGAGGATATTCTGACTGGGT	2535
Qy	2355	TCAAGATGCACTGCGCGGCTGGCGCTCCGTGTACTGCAATGCCAGACGGCGCGCTTCA	2414
Db	2536	TTAAATGATGCAAGGGGCTGGCAATCAATCTACTGCGATGCCACACGACCTTGTTTCA	2595
Qy	2415	AGGGTTCGGCGCGGATCAATCTATCGAACCGTCTCAACAGGTGCTCCGTTGGCGCTGG	2474
Db	2596	AGGGTTCGCAACCAATCAATCTTCCGATCGTCTTAATCAGGTGCTCCGTTGGGCTCTTG	2655
Qy	2475	GGTCCGTCGAGATCTTCTTCAGCGGCGACAGCCCCCTGCTGTACGGCTACAGAAACGGCA	2534
Db	2656	GGTCAGTGGAAATTTCTGCTTAGTAGACATTTGCTCTATCTGCTATGTTTACAAT--GGAC	2712
Qy	2535	ACCTCAAGTGGCTGGAGCGCTTCGCCCTACATCAACACCAACCTACCCCTTCACCTCGC	2594
Db	2713	GATTTGAAGCTTTTGGAGAGCTGGCTTACATCAACACTATTGTATATCCAAATCACATCCA	2772
Qy	2595	TCCCGCTGCTCGCCTACTGCACCCCTCCCGCGCTGCTGCCCTCTCACGGCAAGTTTCATCA	2654
Db	2773	TTCCGCTTATTGCTATTGTGTGTTCCCGCTATCTGCCTCTTACCAATAAATTTATCA	2832
Qy	2655	TGCGGTGATTAGCAGTTCCGACAGCCTCTTCTTCATCGGCCCTTTCATGTCATCTCTCG	2714
Db	2833	TTCTTGAGATTAGCAATTTATGCTGGGATGTTCTTCAATCTTCTTTTCGCTCCATTTTTG	2892
Qy	2715	CGACGGSCATCTCGAGATCGGTGGAGCGGGTGGAGCATCGAGGAGTGGTGAGGAACG	2774
Db	2893	CCACTGGTATTGGAGCTTAGATGGAGTGGTTGGCATTTGAAGATTGGTGAGAAATG	2952
Qy	2775	AGCAGTTCTGGGTCACTCGGCGGGGTGTCGCGCATCTCTTCGCGCGTGTGCAGGGGCTGC	2834

Search completed: March 12, 2005, 00:02:03
Job time : 425 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 21:58:45 ; Search time 1220 Seconds
(without alignments)
16787.522 Million cell updates/sec

Title: US-10-627-132-29
Perfect score: 3443
Sequence: 1 ctgcgtcgccctgcctgcga.....aaaaaaaaaaaaaaaaaaaaa 3443

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3443	100.0	3443	17	US-10-627-132-29
2	3028	87.9	3028	16	US-10-209-059-29
3	3008.6	87.4	3589	17	US-10-425-114-24930
4	2692.8	78.2	3448	18	US-10-437-963-15093
5	2633.2	76.5	2687	17	US-10-425-114-833
6	2495.8	72.5	2830	18	US-10-209-059-25
7	1563.4	45.4	3470	16	US-10-209-059-25
8	1563.4	45.4	3470	17	US-10-627-132-25
9	1523.2	44.2	3517	9	US-09-900-237-13
10	1439.6	41.8	3493	18	US-10-437-963-83811
11	1397.6	40.6	3746	16	US-10-160-719-29
					Sequence 29, Appl
					Sequence 29, Appl
					Sequence 24930, A
					Sequence 15093, A
					Sequence 833, App
					Sequence 139789,
					Sequence 25, Appl
					Sequence 25, Appl
					Sequence 13, Appl
					Sequence 83811, A
					Sequence 29, Appl

12	1397.6	40.6	3746	16	US-10-160-719-49	Sequence 49, Appl
13	1397.6	40.6	3773	16	US-10-209-059-9	Sequence 9, Appl
14	1397.6	40.6	3773	16	US-10-160-719-5	Sequence 5, Appl
15	1397.6	40.6	3773	17	US-10-627-132-9	Sequence 9, Appl
16	1382.4	40.2	3854	18	US-10-425-115-102948	Sequence 102948,
17	1368.6	39.8	3851	17	US-10-393-840-69	Sequence 69, Appl
18	1366.4	39.7	3987	18	US-10-425-115-45192	Sequence 45192, A
19	1364.8	39.6	3786	9	US-09-900-237-7	Sequence 7, Appl
20	1364.8	39.6	3813	16	US-10-209-059-17	Sequence 17, Appl
21	1364.8	39.6	3813	16	US-10-160-719-25	Sequence 25, Appl
22	1364.8	39.6	3813	16	US-10-160-719-45	Sequence 45, Appl
23	1364.8	39.6	3813	17	US-10-627-132-17	Sequence 17, Appl
24	1363.2	39.6	3814	18	US-10-425-115-45185	Sequence 45185, A
25	1325	38.5	3812	18	US-10-425-115-101245	Sequence 101245,
26	1323.4	38.4	3799	16	US-10-267-459-5	Sequence 5, Appl
27	1323.4	38.4	3799	16	US-10-209-059-21	Sequence 21, Appl
28	1323.4	38.4	3799	17	US-10-627-132-21	Sequence 21, Appl
29	1315	38.2	2684	18	US-10-425-115-137279	Sequence 137279
30	1268	36.8	3776	9	US-09-900-237-3	Sequence 3, Appl
31	1266.4	36.8	3505	17	US-10-425-114-31130	Sequence 31130, A
32	1266.4	36.8	3783	18	US-10-425-115-86998	Sequence 86998, A
33	1245.2	36.2	3626	9	US-09-900-237-29	Sequence 29, Appl
34	1242.4	36.1	3962	18	US-10-437-963-72402	Sequence 72402, A
35	1233.2	35.8	3286	18	US-10-425-115-43488	Sequence 43488, A
36	1230.2	35.7	2758	17	US-10-424-599-125586	Sequence 125586,
37	1222.4	35.5	3568	16	US-10-209-059-41	Sequence 41, Appl
38	1222.4	35.5	3568	16	US-10-160-719-1	Sequence 1, Appl
39	1222.4	35.5	3568	17	US-10-627-132-41	Sequence 41, Appl
40	1211.6	35.2	3744	18	US-10-425-115-100413	Sequence 100413,
41	1208.4	35.1	3704	16	US-10-209-059-13	Sequence 13, Appl
42	1208.4	35.1	3704	16	US-10-160-719-57	Sequence 57, Appl
43	1208.4	35.1	3704	17	US-10-627-132-13	Sequence 13, Appl
44	1204.6	35.0	3936	9	US-09-900-237-9	Sequence 9, Appl
45	1204.6	35.0	3969	16	US-10-209-059-49	Sequence 49, Appl

ALIGNMENTS

RESULT 1
US-10-627-132-29
; Sequence 29, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 3443
; TYPE: DNA
; ORGANISM: Zea mays
US-10-627-132-29

Query Match 100.0%; Score 3443; DB 17; Length 3443;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 CTGCGTCGCGCTGCCTGCATCGGAATCTGTGAGCACCTGAGGGTCCGAGCCGAG 60
|||||

Db 1 CTGCGTCGCCCTCGCTCGCAATCGGNAATCTGTGAGACACTGAGGGTCCGAGGCCGAG 60
Qy 61 AGCTAGCTAGCACGCGCGCTCCGGCGCGATGAGAGCCAGCGCGGGCTGCTGCGCGG 120
Db 61 AGCTAGCTAGCACGCGCGCTCCGGCGCGATGAGAGCCAGCGCGGGTGTGCGCGG 120
Qy 121 CTGCGCAACCGGAACGAGTGTGTGTGATCCGGGGCACGAGAACCCCAAGCCGTGCG 180
Db 121 CTGCGCAACCGGAACGAGTGTGTGTGATCCGGGGCACGAGAACCCCAAGCCGTGCG 180
Qy 181 GGGCTGAGCGGCGAGTGTGCGAGATATGCGGCGACGAGTGGGCTCACGGTGCAGCG 240
Db 181 GGGCTGAGCGGCGAGTGTGCGAGATATGCGGCGACGAGTGGGCTCACGGTGCAGCG 240
Qy 241 CGACCTCTTCTGTCGCTCGCAACGAGTGGGCTTCCCGTGTGCGGCGCTGTACAGGTA 300
Db 241 CGACCTCTTCTGTCGCTCGCAACGAGTGGGCTTCCCGTGTGCGGCGCTGTACAGGTA 300
Qy 301 CGAGCGCGGGAGGGCACGAGAACTGCCCCAGTGCAGAGCGGCTAACAGCGCTCAA 360
Db 301 CGAGCGCGGGAGGGCACGAGAACTGCCCCAGTGCAGAGCGGCTAACAGCGCTCAA 360
Qy 361 GGGGAGCCGAGGGTTGCCGGGACGATGACGAGGAGGACATCGACGACCTGGAGCACGA 420
Db 361 GGGGAGCCGAGGGTTGCCGGGACGATGACGAGGAGGACATCGACGACCTGGAGCACGA 420
Qy 421 GTTCAAATCGACGACGAGAACTCAGCAGAGGAGCTGAGGGCAACATGACAGAACAGCCA 480
Db 421 GTTCAAATCGACGACGAGAACTCAGCAGAGGAGCTGAGGGCAACATGACAGAACAGCCA 480
Qy 481 GATCACCGAGCGATGTCACGCGAGATGAGCTACGGAGGGGCGCGACGAGCGGA 540
Db 481 GATCACCGAGCGATGTCACGCGAGATGAGCTACGGAGGGGCGCGCGACGAGCGGA 540
Qy 541 CGSCAACACACCCCGAGATCCCGCCCATCATCACCGCTCCGCTCCGTCGCGGTGAG 600
Db 541 CGSCAACACACCCCGAGATCCCGCCCATCATCACCGCTCCGCTCCGTCGCGGTGAG 600
Qy 601 CGGTGAGTTTCCGATTTACAAACGGGTATGGCCACGGCGAGTCTCGTCTTCCCTGCACAA 660
Db 601 CGGTGAGTTTCCGATTTACAAACGGGTATGGCCACGGCGAGTCTCGTCTTCCCTGCACAA 660
Qy 661 GGGCATCATCGGTACCTGTGTCTGAGCGAGGAGTGCCAAAGTGGGACGAGAAAGA 720
Db 661 GGGCATCATCGGTACCTGTGTCTGAGCGAGGAGTGCCAAAGTGGGACGAGAAAGA 720
Qy 721 AGTGAGCTGGAAGGAGGATGACGACTGGAAGTCCAGCAGGGCATCTCGCGCGCGG 780
Db 721 AGTGAGCTGGAAGGAGGATGACGACTGGAAGTCCAGCAGGGCATCTCGCGCGCGG 780
Qy 781 CGCCGATCCCGAAGACATGACGCCGACGCTGGCACTGAAACGAGCGAGGCGAGCCGCT 840
Db 781 CGCCGATCCCGAAGACATGACGCCGACGCTGGCACTGAAACGAGCGAGGCGAGCCGCT 840
Qy 841 GTGCGAGAGGTGTCATCGCTGAGAGAAAGTGAACCGGTACCGGATGGTGAATCGTGGT 900
Db 841 GTGCGAGAGAGGTGTCATCGCTGAGAGAAAGTGAACCGGTACCGGATGGTGAATCGTGGT 900
Qy 901 GGGTCTGTGCTCGCTTCTTCTCCGTTACCGTATCTTGTGACCCCGTCCCGGACGC 960
Db 901 GGGTCTGTGCTCGCTTCTTCTCCGTTACCGTATCTTGTGACCCCGTCCCGGACGC 960
Qy 961 CATCGGGCTGTGGCTCGTCTCATCATCTGCGAGATCTGGTTCGCCATCTCTGTGATCCT 1020
Db 961 CATCGGGCTGTGGCTCGTCTCATCATCTGCGAGATCTGGTTCGCCATCTCTGTGATCCT 1020
Qy 1021 CGACGAGTTCCCAAGTGGTTCCCAATCGACCGCGAGAGTACCTCGACCGCTCTCCCT 1080
Db 1021 CGACGAGTTCCCAAGTGGTTCCCAATCGACCGCGAGAGTACCTCGACCGCTCTCCCT 1080
Qy 1081 CAGGTACGAGAGGAAGGGAGCCGCTCGTGTGCGGGTGGAATGCTGTGTGAGCAC 1140
Db 1081 CAGGTACGAGAGGAAGGGAGCCGCTCGTGTGCGGGTGGAATGCTGTGTGAGCAC 1140

Qy 1141 GGTGGAACCGCTCAAGGAGCGCGCTGTGTGACCGGCAACACCGTGTCTTCCATCTCTCGC 1200
Db 1141 GGTGGAACCGCTCAAGGAGCGCGCTGTGTGACCGGCAACACCGTGTCTTCCATCTCTCGC 1200
Qy 1201 CGTAGACTACCCCGTGGACAAGGTCTCTGTCTGCTAGCTTCGACGAGCGCGCTGATGCT 1260
Db 1201 CGTAGACTACCCCGTGGACAAGGTCTCTGTCTGCTAGCTTCGACGAGCGCGCTGATGCT 1260
Qy 1261 GAGCTTCGAGTCTGCTGCGAGAGCGCGAGTTTCGCGCGCAAGTGGTGCCCTTCTGCAA 1320
Db 1261 GAGCTTCGAGTCTGCTGCGAGAGCGCGAGTTTCGCGCGCAAGTGGTGCCCTTCTGCAA 1320
Qy 1321 GAAGTTTCGACATCGAGCCCGCGCGCTGGAGTTTCTACTTCTCGCTCAAGTGCAGTACCT 1380
Db 1321 GAAGTTTCGACATCGAGCCCGCGCGCTGGAGTTTCTACTTCTCGCTCAAGTGCAGTACCT 1380
Qy 1381 CAAGGACAAGGTGACGCCCACTTCTGTGACGAGCGCGCGCCCATGAAGAGAGATGA 1440
Db 1381 CAAGGACAAGGTGACGCCCACTTCTGTGACGAGCGCGCGCCCATGAAGAGAGATGA 1440
Qy 1441 GGAGTTCAAGGTCCGGATCAACGCGCTGTGTGGCCAAAGGCCATGAAGTGCAGAGGG 1500
Db 1441 GGAGTTCAAGGTCCGGATCAACGCGCTGTGTGGCCAAAGGCCATGAAGTGCAGAGGG 1500
Qy 1501 GTGGATCATGAAGGACGGCACGCGCTGGCCCGGGAAACAACACCCGCGACACCCCGGCAT 1560
Db 1501 GTGGATCATGAAGGACGGCACGCGCTGGCCCGGGAAACAACACCCGCGACACCCCGGCAT 1560
Qy 1561 GATCAGGTGTCTCGGCGCACGCGCGGCGCACGACGAGGGCAACGAGCTGCCCG 1620
Db 1561 GATCAGGTGTCTCGGCGCACGCGCGGCGCACGACGAGGGCAACGAGCTGCCCG 1620
Qy 1621 CCTCGTACGCTCTCCGCTGAGAGCGCGCGGATTCACAGCACCAAGAGGCCGCGC 1680
Db 1621 CCTCGTACGCTCTCCGCTGAGAGCGCGCGGATTCACAGCACCAAGAGGCCGCGC 1680
Qy 1681 CATGAACGCTCTGATTCGGGTCTCTCCGCTGTGTGACCAACCGCCCATTCATGTCAACTT 1740
Db 1681 CATGAACGCTCTGATTCGGGTCTCTCCGCTGTGTGACCAACCGCCCATTCATGTCAACTT 1740
Qy 1741 GGACTGTGATCACTACATCAACAGAGGCGCATCCGGGAGGCCATGTGTCTCTCAT 1800
Db 1741 GGACTGTGATCACTACATCAACAGAGGCGCATCCGGGAGGCCATGTGTCTCTCAT 1800
Qy 1801 GGAACCTCAGGTCCGGCGGAAGGTCTGCTACGTTTCAGTTCCCGCAGAGTTTCAAGGCAT 1860
Db 1801 GGAACCTCAGGTCCGGCGGAAGGTCTGCTACGTTTCAGTTCCCGCAGAGTTTCAAGGCAT 1860
Qy 1861 CGACGTGACGACCGCATACGCTAAACAGGAACACCGTCTTTCGACATCAACATGAAGG 1920
Db 1861 CGACGTGACGACCGCATACGCTAAACAGGAACACCGTCTTTCGACATCAACATGAAGG 1920
Qy 1921 GCTGAGCGGATCCAGGCGCGGTGTACGTTCGGGACAGGGTGGTGTTCGGGCGCGAGC 1980
Db 1921 GCTGAGCGGATCCAGGCGCGGTGTACGTTCGGGACAGGGTGGTGTTCGGGCGCGAGC 1980
Qy 1981 GCTCTACGCTACAAACCTCCCAAGGACCAAGAGGCCCAAGATGGTGAACCTGCCACTG 2040
Db 1981 GCTCTACGCTACAAACCTCCCAAGGACCAAGAGGCCCAAGATGGTGAACCTGCCACTG 2040
Qy 2041 CTGCGCGTCTTCGGCGCAAGAGCGGAACAACGCAAGGAGCGGCTCCCGAGGGCAC 2100
Db 2041 CTGCGCGTCTTCGGCGCAAGAGCGGAACAACGCAAGGAGCGGCTCCCGAGGGCAC 2100
Qy 2101 CGCTGATATGGAGATAGAGCAAGGAGATGCTCATGTCCCATGATGAACCTTCGAGAA 2160
Db 2101 CGCTGATATGGAGATAGAGCAAGGAGATGCTCATGTCCCATGATGAACCTTCGAGAA 2160
Qy 2161 GCGGTTTCGGGCGAGTCCGCGCGTTTCGTACGCTCGAGCGCTGATGGAGGAGCGCGCTCC 2220
Db 2161 GCGGTTTCGGGCGAGTCCGCGCGTTTCGTACGCTCGAGCGCTGATGGAGGAGCGCGCTCC 2220

```
Qy 2221 TCCTTCGTCGAGCCCGCGCGCTCCTCAAGGAGGCCATCATGTCTATCAGCTCGCGGCTA 2280
Db 2221 TCCTTCGTCGAGCCCGCGCGCTCCTCAAGGAGGCCATCATGTCTATCAGCTCGCGGCTA 2280
Qy 2281 CGAGGACAAGACCGACTGGGGCTGGAGCTGGGTGGATCTACGGGTGCGATCAGCGAGGA 2340
Db 2281 CGAGGACAAGACCGACTGGGGCTGGAGCTGGGTGGATCTACGGGTGCGATCAGCGAGGA 2340
Qy 2341 CATCTGACGGGGTTCAAGATGCACTGCGCGGGTGGCGTCCGTGTACTGCAATGCCGAA 2400
Db 2341 CATCTGACGGGGTTCAAGATGCACTGCGCGGGTGGCGTCCGTGTACTGCAATGCCGAA 2400
Qy 2401 GCGGGCGGCTTCAAGGGTCCGGCCGATCAATCTATCGAGCCGTCTCAACAGGTGCT 2460
Db 2401 GCGGGCGGCTTCAAGGGTCCGGCCGATCAATCTATCGAGCCGTCTCAACAGGTGCT 2460
Qy 2461 CCGGTGGCGCTGGGGTCCGTCGAGATCTTCTTCAGCGCGGACAGCCCTGCTGTACGG 2520
Db 2461 CCGGTGGCGCTGGGGTCCGTCGAGATCTTCTTCAGCGCGGACAGCCCTGCTGTACGG 2520
Qy 2521 CTACAAGACGGCAACTCAAGTGGCTGGAGCGCTTCGCCTACATCAACACCACCATCTA 2580
Db 2521 CTACAAGACGGCAACTCAAGTGGCTGGAGCGCTTCGCCTACATCAACACCACCATCTA 2580
Qy 2581 CCCCTTCACCTCGCTCCGGTCTGCGCTACTGCAACCTCCCGCGCTGCTCCTCTAC 2640
Db 2581 CCCCTTCACCTCGCTCCGGTCTGCGCTACTGCAACCTCCCGCGCTGCTCCTCTAC 2640
Qy 2641 CGGCAAGTTCATCATGCGCTGATAGCATGTTCCGACGCTTCGCGAGCTCTTCTTCATCGCCCTT 2700
Db 2641 CGGCAAGTTCATCATGCGCTGATAGCATGTTCCGACGCTTCGCGAGCTCTTCTTCATCGCCCTT 2700
Qy 2701 CATGTCCATCTTCGCGAGCGGCACTCTGGAGATGCGGTGGAGCGGGGTGAGCATCGAGGA 2760
Db 2701 CATGTCCATCTTCGCGAGCGGCACTCTGGAGATGCGGTGGAGCGGGGTGAGCATCGAGGA 2760
Qy 2761 GTGGTGGAGGACGAGCAGTCTTGGGTCTATGCGGGGTGTCGCGGATCTCTTCGCGGT 2820
Db 2761 GTGGTGGAGGACGAGCAGTCTTGGGTCTATGCGGGGTGTCGCGGATCTCTTCGCGGT 2820
Qy 2821 CGTGACGGGCTGTCTCAAGTCTCTCGCGGATGACACCAACTTCAACGCTCACTCCAA 2880
Db 2821 CGTGACGGGCTGTCTCAAGTCTCTCGCGGATGACACCAACTTCAACGCTCACTCCAA 2880
Qy 2881 GGCACCGGCGACGAGGACGAGTTCGCGGAGCTCTACGCTTCAAGTGGACCAAGCT 2940
Db 2881 GGCACCGGCGACGAGGACGAGTTCGCGGAGCTCTACGCTTCAAGTGGACCAAGCT 2940
Qy 2941 CCTCATCCGCGCCACGACGCTGCTCATCATTAAGCTCATCGGGTCTGTCGCGGATCTC 3000
Db 2941 CCTCATCCGCGCCACGACGCTGCTCATCATTAAGCTCATCGGGTCTGTCGCGGATCTC 3000
Qy 3001 CGACGCTCATCAACAAAGGTTACAGTCTTGGGGGCGCTCTTGGCAAGCTCTTCTTCG 3060
Db 3001 CGACGCTCATCAACAAAGGTTACAGTCTTGGGGGCGCTCTTGGCAAGCTCTTCTTCG 3060
Qy 3061 CTTCTGGGTCTCATCGTCCACCTCTACCGGTTCTCAAGGGGCTCATGGGGGCGCGAGACAG 3120
Db 3061 CTTCTGGGTCTCATCGTCCACCTCTACCGGTTCTCAAGGGGCTCATGGGGGCGCGAGACAG 3120
Qy 3121 GAGCCGACCGTGTGTGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3180
Db 3121 GAGCCGACCGTGTGTGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3180
Qy 3181 GGTGAGATCGACCTTTTCAAGTCTGTCAGGACCAAGGCGCGGAGTCAAGGAGTGTGCGAT 3240
Db 3181 GGTGAGATCGACCTTTTCAAGTCTGTCAGGACCAAGGCGCGGAGTCAAGGAGTGTGCGAT 3240
Qy 3241 CAATTGCTGAGCTGTTTATTAAGGTTCAAAATTTCTGGAGTCTTGTGATAGGAGAAAAA 3300
Db 3241 CAATTGCTGAGCTGTTTATTAAGGTTCAAAATTTCTGGAGTCTTGTGATAGGAGAAAAA 3300
Qy 3301 ACAATTTAGAAATTTTGTAAAGTTTGTGTCTGTGTAAATGTATGTTACCCAGAAATTTGTCG 3360
```

```
Db 3301 ACAATTTAGAAATTTTGTAAAGTTTGTGTCTGTGTAAATGTATGTTACCCAGAAATTTGTCG 3360
Qy 3361 GACGAGGAAATTTGAACAAAGGACAGGTTTGTATTGTTAAATGGCAAAAAAAGAAAAA 3420
Db 3361 GACGAGGAAATTTGAACAAAGGACAGGTTTGTATTGTTAAATGGCAAAAAAAGAAAAA 3420
Qy 3421 AAAAAAAGAAAAAAGAAAAAAGAAAAA 3443
Db 3421 AAAAAAAGAAAAAAGAAAAAAGAAAAA 3443

RESULT 2
US-10-209-059-29
; Sequence 29, Application US/10209059
; Publication No. US2003016383A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 3028
; TYPE: DNA
; ORGANISM: Zea mays
US-10-209-059-29

Query Match 87.9%; Score 3028; DB 16; Length 3028;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3028; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 416 CACGAGTTCAACATCGACGACGAGAAATCAGCAGGCGAGCTGGAGGGCAATCAGAGAAC 475
Db 1 CACGAGTTCAACATCGACGACGAGAAATCAGCAGGCGAGCTGGAGGGCAATCAGAGAAC 60
Qy 476 AGCCAGATCACCAGGCGATGCTCGCGGAGATGAGCTACGAGGGGCGCGAGCAG 535
Db 61 AGCCAGATCACCAGGCGATGCTCGCGGAGATGAGCTACGAGGGGCGCGAGCAG 120
Qy 536 GCGACGCGCAACACACCCCGCAGATCCCGCCCATCATCACCGGCTCCCGTCCGTCGCG 595
Db 121 GCGACGCGCAACACACCCCGCAGATCCCGCCCATCATCACCGGCTCCCGTCCGTCGCG 180
Qy 596 GTGAGCGGTGAGTTTCCGATTACCAACGGGTATGCGCAACGGGCGAGTCTCGTCTTCCCTG 655
Db 181 GTGAGCGGTGAGTTTCCGATTACCAACGGGTATGCGCAACGGGCGAGTCTCGTCTTCCCTG 240
Qy 656 CACAAGCGCATCCATCGGTACCTGCTGTGAGCAGGAGTCCAGTGGGACGAGAG 715
Db 241 CACAAGCGCATCCATCGGTACCTGCTGTGAGCAGGAGTCCAGTGGGACGAGAG 300
Qy 716 AAAAAAGTGAAGGAGGAGGAGTGGACGACTGGAAGTCCAAAGCAGGAGCATCTCTCGGC 775
Db 301 AAAAAAGTGAAGGAGGAGGAGTGGACGACTGGAAGTCCAAAGCAGGAGCATCTCTCGGC 360
Qy 776 GCGGCGCGCATCCCGAAGACATGAGCGCGAGCTGGCACTGAAACGAGGCGAGCAG 835
Db 361 GCGGCGCGCATCCCGAAGACATGAGCGCGAGCTGGCACTGAAACGAGGCGAGCAG 420
Qy 836 CCGCTCTCGAGGAGGAGTGTGATCGGTGAGCAAGGTGAAACCGGTACCGGATGGTATC 895
Db 421 CCGCTCTCGAGGAGGAGTGTGATCGGTGAGCAAGGTGAAACCGGTACCGGATGGTATC 480
```

Qy	896	GTGTTGGCTCTCGTTGTGCTCGCCTTCTTCTCTCGGTACCGTATCTGTGCAACCCGCTCCCG	955
Db	481	GTGTTGGCTCTCGTTGTGCTCGCCTTCTTCTCTCGGTACCGTATCTGTGCAACCCGCTCCCG	540
Qy	956	GACGCCATCGGCTGTGGCTGCTCATCATCTGCGAGATCTGGTTCCGGCATCTCCTGG	1015
Db	541	GACGCCATCGGCTGTGGCTGCTCTCCATCATCTGCGAGATCTGGTTTCGTCATCTCCTGG	600
Qy	1016	ATCTCTGACAGTTTCCCAAGTGGTTCCCATCGACCGGAGAGTACTCTCGACCGCCTC	1075
Db	601	ATCTCTGACAGTTTCCCAAGTGGTTCCCATCGACCGGAGAGTACTCTCGACCGCCTC	660
Qy	1076	TCCCTCAGGTACGAGAGGGAAGGGAGCCGTCGCTGCTGTCTCGCGGTGGACCTGTTCGTG	1135
Db	661	TCCCTCAGGTACGAGAGGGAAGGGAGCCGTCGCTGCTGTCTCGCGGTGGACCTGTTCGTG	720
Qy	1136	AGCACGTTGACCCGCTCAAGAGAGCCCGCTGTGTGACCGGCAACACCGTGTCTCTCATC	1195
Db	721	AGCACGTTGACCCGCTCAAGAGAGCCCGCTGTGTGACCGGCAACACCGTGTCTCTCATC	780
Qy	1196	CTGCGCTGTAGACTACCCCGTGGACAAGGTCTCTGTCTACGTCTCTCGACGAGCGCGGTG	1255
Db	781	CTGCGCTGTAGACTACCCCGTGGACAAGGTCTCTGTCTACGTCTCTCGACGAGCGCGGTG	840
Qy	1256	ATGCTGACGTTTTCAGTTCGCTGTGCGAGACGGCCGAGTTCGCGCGCAAGTGGGTGCCCTTC	1315
Db	841	ATGCTGACGTTTTCAGTTCGCTGTGCGAGACGGCCGAGTTCGCGCGCAAGTGGGTGCCCTTC	900
Qy	1316	TGCAAGAGTTTCGCGCATCGAGCCCGCCGCGCTTCGCTCTCAAGTTCGCTCAAGTTCGAC	1375
Db	901	TGCAAGAGTTTCGCGCATCGAGCCCGCCGCGCTTCGCTCTCAAGTTCGCTCAAGTTCGAC	960
Qy	1376	TACCTCAAGNACAGTTCGACCCACCTTTCGTGACGAGAGCGCGCCCATGAAGAGAGAG	1435
Db	961	TACCTCAAGNACAGTTCGACCCACCTTTCGTGACGAGAGCGCGCCCATGAAGAGAGAG	1020
Qy	1436	TATGAGGAGTTTCAAGTTCGATCAACGCTGTGTGCGCAAGGCCATGAAGTTCGCGGCA	1495
Db	1021	TATGAGGAGTTTCAAGTTCGATCAACGCTGTGTGCGCAAGGCCATGAAGTTCGCGGCA	1080
Qy	1496	GAGGGTGGATTCATGAAGGACGGCAACGCGTGGCCCGGGAACAAACCCCGCACACCCC	1555
Db	1081	GAGGGTGGATTCATGAAGGACGGCAACGCGTGGCCCGGGAACAAACCCCGCACACCCC	1140
Qy	1556	GGCATGATCCAGGTGTTCTTGGGCCACAGCGGGGCCACGACACCGGAGGCGACGAGCTG	1615
Db	1141	GGCATGATCCAGGTGTTCTTGGGCCACAGCGGGGCCACGACACCGGAGGCGACGAGCTG	1200
Qy	1616	CCCGCTCTGTGACGTTCTCCGTGAGAAGCGCCCGGGATTCACAGCACCAAGAGAGGCC	1675
Db	1201	CCCGCTCTGTGACGTTCTCCGTGAGAAGCGCCCGGGATTCACAGCACCAAGAGAGGCC	1260
Qy	1676	GGCGCCATGAACGCTCTGATTCGCGTCTCCGCGTGTGACCAACGCGCATTCATGCTC	1735
Db	1261	GGCGCCATGAACGCTCTGATTCGCGTCTCCGCGTGTGACCAACGCGCATTCATGCTC	1320
Qy	1736	AACTTGACATGTGATCACTACATCAACAGCAAGGCCATTCGCGGAGGCCATGTGCTTC	1795
Db	1321	AACTTGACATGTGATCACTACATCAACAGCAAGGCCATTCGCGGAGGCCATGTGCTTC	1380
Qy	1796	CTCATGACCCCTCAGGTTCGCGCGGAAGGTCTGCTACGTTTCAGTTTCGCGAGAGTTTCGAC	1855
Db	1381	CTCATGACCCCTCAGGTTCGCGCGGAAGGTCTGCTACGTTTCAGTTTCGCGAGAGTTTCGAC	1440
Qy	1856	GGCATCGACGTGACGACCGGATACGCTAAACAGGAACACCGCTTCTTTCGACATCAACATG	1915
Db	1441	GGCATCGACGTGACGACCGGATACGCTAAACAGGAACACCGCTTCTTTCGACATCAACATG	1500
Qy	1916	AAGGGGCTGACCGGCATCCAAAGGCCCGGTGTACGTTCGGGACAGGGTTCGTGTTCCGGCGC	1975
Db	1501	AAGGGGCTGACCGGCATCCAAAGGCCCGGTGTACGTTCGGGACAGGGTTCGTGTTCCGGCGC	1560

1976	Qy	CAGCGCGTCTACGGCTACAAACCTCCCAAGGGGACCCAAAGAGGCCAAAGATGGTGACTGCG	2035
1561	Db	CAGCGCGTCTACGGCTACAAACCTCCCAAGGGGACCCAAAGAGGCCAAAGATGGTGACTGCG	1620
2036	Qy	GACTGTCGCCGCTTCGCGCCGCAAGAGCGGAAACAGCCAAAGAGCGGCTGCCGGAG	2095
1621	Db	GACTGTCGCCGCTTCGCGCCGCAAGAGCGGAAACAGCCAAAGAGCGGCTGCCGGAG	1680
2096	Qy	GGCACCGCTGATATGGGAGTAGTAGCGACAAGGAGATGCTCATGTGCCACATGAATCTTC	2155
1681	Db	GGCACCGCTGATATGGGAGTAGTAGCGACAAGGAGATGCTCATGTGCCACATGAATCTTC	1740
2156	Qy	GAGAAGCGGTTCCGGCAGTCCGCGCGTTCGTCACGTCGACGTGTGATGAGGAGGAAGCGCGC	2215
1741	Db	GAGAAGCGGTTCCGGCAGTCCGCGCGTTCGTCACGTCGACGTGTGATGAGGAGGAAGCGCGC	1800
2216	Qy	GTCCCTCTTCGTCGAGCCCGCGCGCTCCTCAAGGAGGCGCATCCATGTCATCAGCTGC	2275
1801	Db	GTCCCTCTTCGTCGAGCCCGCGCGCTCCTCAAGGAGGCGCATCCATGTCATCAGCTGC	1860
2276	Qy	GGCTACGAGGACAAGACCCGACTGGGGCTGGAGCTGGGGTGAGTCTACGGGTGATCAGC	2335
1861	Db	GGCTACGAGGACAAGACCCGACTGGGGCTGGAGCTGGGGTGAGTCTACGGGTGATCAGC	1920
2336	Qy	GAGGACATCCTGACGGGGTTCAAGATGCACCTGCCGGGGTGGCGCTCCGTGTACTGCA	2395
1921	Db	GAGGACATCCTGACGGGGTTCAAGATGCACCTGCCGGGGTGGCGCTCCGTGTACTGCA	1980
2396	Qy	CCGAGCGGCGGGGTTCAAGGGGTTCGCGCCCGATCAATCTATCGAGCCGTCTCAAC	2455
1981	Db	CCGAGCGGCGGGGTTCAAGGGGTTCGCGCCCGATCAATCTATCGAGCCGTCTCAAC	2040
2456	Qy	GTGCTCGGTGGCGCTGGGGTTCGTCGAGATCTTTTCAGCGGGCACAGCCCGCTGCTG	2515
2041	Db	GTGCTCGGTGGCGCTGGGGTTCGTCGAGATCTTTTCAGCGGGCACAGCCCGCTGCTG	2100
2516	Qy	TACGGCTACAAGAACCGGCAACCTCAAGTGGCTGGAGCGCTTCGCCTACATCAACAC	2575
2101	Db	TACGGCTACAAGAACCGGCAACCTCAAGTGGCTGGAGCGCTTCGCCTACATCAACAC	2160
2576	Qy	ATCTACCCCTTACGCTCGCTCCGCTGCTCGCTACTGACACCTCCCGCGGCTGCTGC	2635
2161	Db	ATCTACCCCTTACGCTCGCTCCGCTGCTCGCTACTGACACCTCCCGCGGCTGCTGC	2220
2636	Qy	CTACCGGCAAGTTTCATCATGCGTGCATTAGACAGTTTCGCCAGCGCTTCTTCATCG	2695
2221	Db	CTACCGGCAAGTTTCATCATGCGTGCATTAGACAGTTTCGCCAGCGCTTCTTCATCG	2280
2696	Qy	CTCTTCATGTCCATCTTCGCGACGGGCATCCTGGAGATGCGGTGGAGCGGGTGAG	2755
2281	Db	CTCTTCATGTCCATCTTCGCGACGGGCATCCTGGAGATGCGGTGGAGCGGGTGAG	2340
2756	Qy	GAGGAGTGTGGAGGAACGAGAGTTCGGGTGATCGGCGGCTGTGTCGGCGCATCTCT	2815
2341	Db	GAGGAGTGTGGAGGAACGAGAGTTCGGGTGATCGGCGGCTGTGTCGGCGCATCTCT	2400
2816	Qy	GCGTCTGTGAGGGCTGCTCAAGGTCCTCGCGGGATCGACACCAACTTCACCGTC	2875
2401	Db	GCGTCTGTGAGGGCTGCTCAAGGTCCTCGCGGGATCGACACCAACTTCACCGTC	2460
2876	Qy	TCCAAAGCCACCGGCGACGAGGACGAGTTCGCCGAGCTCTACGCGCTTCAAGTGA	2935
2461	Db	TCCAAAGCCACCGGCGACGAGGACGAGTTCGCCGAGCTCTACGCGCTTCAAGTGA	2520
2936	Qy	ACGCTCTCATTCGCGCCCAACAGCTGCTCATCATTAACGTCATCGGCGTGTGCGCG	2995
2521	Db	ACGCTCTCATTCGCGCCCAACAGCTGCTCATCATTAACGTCATCGGCGTGTGCGCG	2580
2996	Qy	ATCTCCGACGCCATCAACAGGGTACAGTTCCTGGGGGCCCTCTTCGCGAAGCTCTTC	3055
2581	Db	ATCTCCGACGCCATCAACAGGGTACAGTTCCTGGGGGCCCTCTTCGCGAAGCTCTTC	2640
3056	Qy	TTCCGCTTCTGGGTCATGTCGACCTCTACCCGTTCTCAAGGGGCTCATGGGGCGC	3115

Db	1473	CTGCAAGAGTTCTGCATCGAGCCCGCGCCCGAGTTCCTACTTCTCGTCAAGGTCGA	1532
Qy	1375	CTACTCTCAAGGACAAGGTGCAGGCCACTTCGTGTCAGAGCGCGCGCCCATGAAGAGAGA	1434
Db	1533	TTACTCTCAAGGACAAAGGTGCAGGCCACTTCGTGTCAGAGCGCGCGCCCATGAAGAGAGA	1592
Qy	1435	GTATGAGGAGTTCAAGGTCGCGATCAACCGGCTGGTGGCCAAAGGCCATGAAGGTGCCGCG	1494
Db	1593	GTACGAGGAGTTCAAGGTCCGATCAACCGCTGTTGCCCAAGGCCATGAAGGTGCCGCG	1652
Qy	1495	AGAGGGGTGGATCATATGAAGGACGGCAACCGCTGCGCCCGGGAAACAAACCCCGGACCAACC	1554
Db	1653	AGAGGGGTGGATCATATGAAGGACGGCAACCGCTGCGCCCGGGAAACAAACCCCGGACCAACC	1712
Qy	1555	CGCATCATCCAGGTGTTCTTGCGCACAGCGGGCGGACGACACCGAGGGCACACGAGCT	1614
Db	1713	TGGCATGATCCAGGTGTTCTTGCGCACAGCGGGCGGACGACACCGAGGGCAACGAGCT	1772
Qy	1615	GCCCGGCTCGTGATGTCCTCCGTGAGAAAGCGCCCGGATTCACAGCAACCAAGAAAGGC	1674
Db	1773	GCCCGGCTCGTGATGTCCTCCGTGAGAAAGCGCCCGGATTCACAGCAACCAAGAAAGGC	1832
Qy	1675	CGCGCCCATGAACGCTCTGATTTGCGTCTCCCGCTGCTGACCAACCGGCGCATTCATGCT	1734
Db	1833	CGCGCCCATGAACGCTCTGATTCGCGTCTCCCGCTGCTGACCAACCGGCGCATTCATGCT	1892
Qy	1735	CAACTTGGACTGTGATCACTACATCAACACGCAAGGCCATCCGGAGAGCCATGTGCTT	1794
Db	1893	CAACTTGGACTGTGATCACTACATCAACACGCAAGGCCATCCGGAGAGCCATGTGCTT	1952
Qy	1795	CCTCATGGAACCTCAGGTCCGGCCGGAAGTCTGTACTGTTCAAGTTCGCCGAGAGGTTCTGA	1854
Db	1953	CCTCATGGAACCTCAGGTCCGGCCGGAAGTCTGTACTGTTCAAGTTCGCCGAGAGGTTCTGA	2012
Qy	1855	CGGCATCGACGTGACGACCGATACGCTAACAGGAAACCCGTCTTTTCGACATCAACAT	1914
Db	2013	CGGCATCGACGTGACGACCGATACGCGAAACGGAAACCCGTCTTTTCGACATCAACAT	2072
Qy	1915	GAAGGGCTCGACGGCATCCAGCCCGGTGTACTCGGACAGAGGTGCGTGTCCGGCG	1974
Db	2073	GAAGGGCTCGACGGCATCCAGGGCCCGGTCTACTGTCGGACAGAGGTGCGTGTCCGGCG	2132
Qy	1975	CAAGGCGCTCTACGGCTACAACCTCCCAAGGGACCCCAAGAGGCCCAAGATGTGACCTG	2034
Db	2133	CAAGGCGCTCTACGGCTACAACCTCCCAAGGGCCCAAGAGGCCCAAGATGTGACCTG	2192
Qy	2035	CGACTGTCGCCCGTCTTCGGCCGCAAGAACGGAAACACGCCCAAGGACGGGTCCCGGA	2094
Db	2193	CGACTGTCGCCCGTCTTCGGCCGCAAGAACGGAAACACGCCCAAGGATGGGTCCCGGA	2252
Qy	2095	GGGCACCGCTGATATGGAGTAGATAGGACAGGAGATGCTCATGTCCACATGAACCTT	2154
Db	2253	GGGCACCGCGGATATAT---CGTGGACACGCAACGAGAGATGCTCATGTGCCAGATGAACCTT	2309
Qy	2155	CGAAGCGGTTCCGGCAGTCCGGCGGTTGCTCACGTTCGACGTGATGAGAGAAAGCGG	2214
Db	2310	CGAAGACGGTTCCGGCAGTTCGGCGGCTTCGTCAAGTTCACGCTCATGGAGAGAGCGG	2369
Qy	2215	CGTCCCTCTTTCGTTCGAGCCCGCGCGCTCCTCAAGAGGCCATCCATGTCATCAGCTG	2274
Db	2370	CGTCCCTCTTTCGTTCGAGCCCGCGCTCCTCAAGAGGCCATCCATGTCATCAGCTG	2429
Qy	2275	CGGCTACGAGGACAAAGACCGACTGGGGCTGGAGCTGGGGTGGATCTACCGGTCGATCAC	2334
Db	2430	CGGCTACGAGGACAAAGACCGACTGGGGCTGGAGCTGGGGTGGATCTACCGGTCGATCAC	2489
Qy	2335	GGAGGACATCTCTGACGGGTTCAAGATGCACTCCGCGGGTGGCGCTCCGCTGACTGTCAT	2394
Db	2490	GGAGGACATCTCTGACGGGTTCAAGATGCACTCCGCGGGTGGCGCTCCGCTGACTGTCAT	2549
Qy	2395	GCCGAAGCGGGCGGCTTCAAGGGGTCGGCGCGGATCAATCTATTCGAGACCGCTCTCAACCA	2454

2550	GCCGAA	GCGGGCGGCGGTTCAAGGGGTGCGCGCGATCAACCTGTCCGACGCTCTCAACCA	2609
Qy	Db		
2455	GGTGCT	CCGGTGGGCGCTGGGGTCCGTCGAGATCTTCTTCAGCGGGCACAAGCCCCCTGCT	2514
Qy	Db		
2610	GGTGCT	CCGGTGGGCGCTGGGGTCCGTCGAGATCTTCTTCAGCGGGCACAAGCCCCCTGCT	2669
Qy	Db		
2515	GTACGGCTA	CAAGAACGGCAACTCAAGTGGCTGGAGCGCTTCGCCCTACATCAACACCAAC	2574
Qy	Db		
2670	GTACGGCTA	CAAGAACGGCAACCTCAAGTGGCTGGAGCGCTTCGCCCTACATCAACACCAAC	2729
Qy	Db		
2575	CATCTACCCCTT	CACCTCGCTCCGCTGCTCGCTACTGCAACCTCCGCCGCTCTGCGCTCTGCGCT	2634
Qy	Db		
2730	CATCTACCCCTT	CACCTCGCTCCGCTGCTCGCTACTGTACCTCCGACGCCGCTCTCGCT	2789
Qy	Db		
2635	CCTCACCGGCAAGTT	CATCATATGCCGTTCGATTAGCAGCTTCGCCAGCGCTCTTCTTCATCGC	2694
Qy	Db		
2790	CCTCACCGGCAAGTT	CATCATATGCCGTTCGATTAGCAGCTTCGCCAGCGCTCTTCTTCATCGC	2849
Qy	Db		
2695	CCTCTTCATGTCCAT	TTTCCGAGAGGGCATCTCTGGAGATCGGTGGAGCGGGGTGAGCAT	2754
Qy	Db		
2850	CCCTCTTCATGTCCAT	TTTCCGAGAGGGCATCTCTGGAGATCGGTGGAGCGGGGTGAGCAT	2909
Qy	Db		
2755	CGAGGAGTGT	TGAGGAAACGAGACAGTCTCTGGGTCTCATCGCGCGGTGTCCCGCATCTCTTT	2814
Qy	Db		
2910	CGAGGAGTGT	TGAGGAAACGAGACAGTCTCTGGGTCTCATCGCGCGGTGTCCCGCATCTCTTT	2969
Qy	Db		
2815	CGCGCTCGTGCAGGGCGCTGT	CTCAAGGTCTCTCGCGGGATCGACCAACCTTCACCGCTCAC	2874
Qy	Db		
2970	CGCGCTCGTGCAGGGCGCTGT	CTCAAGGTCTCTCGCGGGATCGACCAACCTTCACCGCTCAC	3029
Qy	Db		
2875	CTCCAGGGCCACCGGGCAGCAGACGACGAGTTCGCCGAGCTCTACGCGCTCTCAAGTGGAC	2934	
Qy	Db		
3030	CTCCAGGGCCACCGGGCAGCAGTGGACGACGAGTTTCGCCGAGCTCTACGCGTCTCAAGTGGAC	3089	
Qy	Db		
2935	CACGCTCCTCATCCCGCCACACCGCTGTCTCATCTTAAGCTCATCGGCGTCTGGCGCGG	2994	
Qy	Db		
3090	CACGCTCCTCATCCCGCCACACCGCTGTCTCATCTTAAGCTCATCGGCGTCTGGCGCGG	3149	
Qy	Db		
2995	CATCTCGAGCGCCATCAACAAACGGGTACAGCTCTGGGGGCCCTCTTCGGCAAGCTCTTT	3054	
Qy	Db		
3150	CATCTCGAGCGCCATCAACAAACGGGTACCAAGTCTGGGGGCCCTCTTCGGCAAGCTCTTT	3209	
Qy	Db		
3055	CTTCGCGCTTCTGGGTCAATCGTTCACCTCTACCCGTTCTCCTCAAGGGGCTCATGGGGCGCCA	3114	
Qy	Db		
3210	CTTCGCGCTTCTGGGTCAATCGTTCACCTCTACCCGTTCTCCTCAAGGGGCTCATGGGGCGCCA	3269	
Qy	Db		
3115	GAACAGGACGCCCAACCGTTGTGTCTCTGGTCCATTTCTGCTGGCGCTCCATCTTCTCCCT	3174	
Qy	Db		
3270	GAACCGGACGCCCAACCATTTGTGTCTCTGGTCCGTTCTGCTGGCGCTCCATCTTCTCCCT	3329	
Qy	Db		
3175	GCTCTGGTTCAGATCGACCCCTTTCTATCGTCAGAACCAAGGGCCCGACGCTCAGGCAAGTG	3234	
Qy	Db		
3330	GCTCTGGTTCAGATCGATCCCTTTCTATCGTCAGAACCAAGGGCCCGACGCTCAGGCAAGTG	3389	
Qy	Db		
3235	TGGCATCAATTCGCTGAGCTGTTTATTAAGGTTCAAAA---TTCTGGAGCTGTGTGCATAGG	3291	
Qy	Db		
3390	TGGCATCAACTGCTGAGCTGCCTATACGCGTTCAAAAAAGTTCTAGAGCTTGTGCAATAGG	3449	
Qy	Db		
3292	GAGAAAAAACAATTTAGAAATTTTGT-----AGGTTGTGTG	3330	
Qy	Db		
3450	GAG-AAAAACACGTTTAGAAATTTTGTGTGTGTATGTTTTGGAGCTCATATGTCATCGTTT	3508	
Qy	Db		
3331	CTGTGAATGTTATGGTACCCAGAAATGTT---CGGACGAGGAATTGAAACAAAGGACAAGG	3386	
Qy	Db		
3509	TGTGTATGTTCTTGGNCCAGAAATGTCTCGGCCGACGAGGAATTGACAAAGGACAAGG	3568	
Qy	Db		
3387	TTTGATT	3393	
Qy	Db		
3569	TTTGAAT	3575	
Qy	Db		

; Sequence 15093, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 15093
; LENGTH: 3448
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_20969C.1
US-10-437-963-15093

Query Match 78.2%; Score 2692.8; DB 18; Length 3448;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 2954; Conservative 0; Mismatches 322; Indels 18; Gaps 5;
QY 67 CCTAGCAGCCGGCTCCGCGCGATGGAGGCCAGCCGGGCTGGTGGCGGCTCGCA 126
DB 3398 GTTTCGGGCTTGTTCGGCGGCATGGAGCGAGCGCGGCTGGTGGCGGCTCGCA 3339
QY 127 CAACCGGAACAGCTGTGTGATCCGGGGCCAGAGGACCCCAAGCCGCTCGCGGGCT 186
DB 3338 CAACCGGAACAGCTGTGTGATCCGGGGCCAGAGGACCCCAAGCCGCTCGCGGGCT 3279
QY 187 GAGCGGGAGGTGTGCGAGATATCGCGGACAGAGTGGGTCTACGGTGGACGGCACCT 246
DB 3278 GAGCGGGAGGTGTGCGAGATATCGCGGACAGAGTGGGTCTACGGTGGACGGCACCT 3219
QY 247 CTTGCTGCTGCAACAGTCCGCTTCCCGTGTGGCGGCTTACGAGTACGAGCG 306
DB 3218 CTTGCTGCTGCAACAGTCCGCTTCCCGTGTGGCGGCTTACGAGTACGAGCG 3159
QY 307 CCGGAGGGACCGCAGAACTCCCGCAGTGAAGACCGCTACAGCGCCTCAAGGGGAG 366
DB 3158 CCGGAGGGACCGCAGAACTCCCGCAGTGAAGACCGCTACAGCGCCTCAAGGGGAG 3099
QY 367 CCGGAGGGTCCCGGGACGATGACGAGGAGGACATCGACCTGGAGACAGATTCAA 426
DB 3098 CCGGAGGGTCCCGGGACGAGGACGAGGAGGACATTCGACCTGGAGACAGATTCAA 3039
QY 427 CATCGACGACGAGATCAGCAGGACGAGTGGAGGCA-----CATGACGACGCA 480
DB 3038 CATCGACGACGAGATCAGCAGGACGAGTGGAGGCA-----CATGACGACGCA 2979
QY 481 GATCAGGAGGATGTGTCAGCGCAGGATGAGTACGGAGGGGCCCCGACGACGCGA 540
DB 2978 CATCAGGAGGATGTGTCAGCGCAGGATGAGTACGGAGGGGCCCCGACGACGCGA 2919
QY 541 CGGCAACAAACCCCGGAGATCCCGCCATCATCACCGGCTCCCGTCCGTCGGGTGAG 600
DB 2918 CGGCAACAGACCCCGC---TCCCGCGGATCATCACCGGCTCCCGTCCGTCGGGTGAG 2862
QY 601 CGGTGAGTTCCGATTACCAACGGGTATGCGCAGCGGAGTCTCGTCTTCCCTGCAAA 660
DB 2861 CGGGAGTTCCCGATTACCAACGGGTATGCGCAGCGGAGTCTCGTCTTCCCTGCAAA 2802
QY 661 GCGCATCCATCCGTACCTGTGTGAGCAGGAGGTGCCAAGTGGGACGAGAAAGA 720
DB 2801 GCGCATCCATCCGTACCTGTGTGAGCAGGAGGTGCCAAGTGGGACGAGAAAGA 2742

QY 721 AGTCAGCTGGAAGGAGGAGGATGGACGATGGAAGTCCAAGCAGGGGCATCTCTCGCGCGG 780
DB 2741 GGTGAGCTGGAAGGAGGAGGATGGACGATGGAAGTCCAAGCAGGGGCATCTCTCGCGCGG 2682
QY 781 CG----CGATCCCGAAGACATGAGCGCGGAGTGGCACTGAACAGAGGGGAGGAGCC 837
DB 2681 CGCCCCGATCCCGAAGACATGAGCGCGGAGTGGCACTGAACAGAGGGGAGGAGCC 2632
QY 838 GCTGTCGAGGAAGGTGTGATCGGTCGAGCAAGGTGAACCGGTACCGGATGGTGTGATCGT 897
DB 2621 GCTGTCGAGGAAGGTGTGATCGGTCGAGCAAGGTGAACCGGTACCGGATGGTGTGATCAT 2562
QY 898 GGTGCGTCTGTTGTCGCTTCTTCTCCCGTACCGTATCTGCAACCCCGTCCCGGA 957
DB 2561 CTTCCGTCGTCGTCGCTTCTTCTCCCGTACCGTATCTGCAACCCCGTCCCGGA 2502
QY 958 GCGCATCGGCTGTGGCTCGTCTCATCTATCTGAGATCTGGTTCGCCATCTCTCTGAT 1017
DB 2501 GCGCATCCGCTGTGGCTCACCTCATCTGAGATCTGGTTCGCCATCTCTCTGAT 2442
QY 1018 CTTGACACAGTTCCCGAAGTGTTCCTCATCTGAGCGGAGACGTACCTCGACCGCTCTC 1077
DB 2441 CTTGACACAGTTCCCGAAGTGTTCCTCATCTGAGCGGAGACGTACCTCGACCGCTCTC 2382
QY 1078 CTTGACGATCGAGGAGGAGGAGCGTCTGCTGTCGCGGTCGACCTGTTCGTGAG 1137
DB 2381 CTTCCGCTACGAGCGCGGAGGAGCGTCTGCTGTCGCGGTCGACCTGTTCGTGAG 2322
QY 1138 CAGGTGAGCCGCTCAAGGAGCGCGCTGTCGACCGCAACACCGTCTCTCATCTCT 1197
DB 2321 CAGGTGAGTCCGCTCAAGGAGCGCGCTGTCGACCGCAACACCGTCTCTCATCTCT 2262
QY 1198 CGCGTGTGATCTACCCCTGAGCAAGTCTCTGTCGTCGACCGCGGCTCGAT 1257
DB 2261 CGCGTGTGATCTACCCCTGAGCAAGTCTCTGTCGTCGACCGCGGCTCGAT 2202
QY 1258 GCTGACGTTTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1317
DB 2201 GCTGACGTTTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2142
QY 1318 CAAGAGTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1377
DB 2141 CAAGAGTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2082
QY 1378 CTTCAAGGACAGAGTGTGAGCCCTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1437
DB 2081 CTTCAAGGACAGAGTGTGAGCCCTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2022
QY 1438 TGAGGATTCGAGTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1497
DB 2021 CGAGGATTCGAGTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1962
QY 1498 GGGTGTGATCATGAAGGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1557
DB 1961 AGGTGTGATCATGAAGGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1902
QY 1558 CATGATCCAGTGTTCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1617
DB 1901 CATGATCCAGTGTTCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1842
QY 1618 CCGCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1677
DB 1841 CCGCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1782
QY 1678 GCGCATGAGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1737
DB 1781 GCGCATGAGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1722
QY 1738 CTTGAGTGTGATCATGATCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1797
DB 1721 CTTGAGTGTGATCATGATCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1662
QY 1798 CATGGACCTTCAGGTTCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1857

Qy	3173	CTGCTCTGGGTCAGGATCGACCCCTTTTCATCTGTCAGGACCAAGGGCCCGAGAGCTCAGGCGAG	3232
Db	2401	CTGCTCTGGGTCAGGATCGACCCCTTTTCATCTGTCAGGACCAAGGGCCCGAGAGCTCAGGCGAG	2460
Qy	3233	TGTGGCATCAATTGCTGAGCTGCTTTATTAAAGGTTCAAATACTCGAGCTTGTGCAATAGGG	3292
Db	2461	TGTGGCATCAATTGCTGAGCTGCTTTATTAAAGGTTCAAATACTCGAGCTTGTGCAATAGGG	2520
Qy	3293	AGAAAAAACAAATTAGAAATTTGCTAAAGGTTGTTGTCTCTAATGTTATGTTACCCAG	3352
Db	2521	AGAAAAAACAAATTAGAAATTTTCTAAGGTTGTTGTCTCTAATGTTATGTTACCCAG	2580
Qy	3353	AATTGTCGACGAGGAATTGAACAAAGGCAAGGTTTGATTGTTTAAATGCGCAAAAAAAA	3412
Db	2581	AATTGTCGACGAGGAATTGAACAAAGGCAAGGTTTGATTGTTTAAATGCGCAATGAGTG	2640
Qy	3413	AAAAAAAAAAAAAAAAAAAAAAAAAAAA 3442	
Db	2641	CAGTCATCTAAATGTAATCCAAATAGCAA 2670	
RESULT 6			
US-10-425-115-139789			
; Sequence 139789, Application US/10425115			
; Publication No. US20040214272A1			
; GENERAL INFORMATION:			
; APPLICANT: La Rosa, Thomas J.			
; APPLICANT: Kovalic, David K.			
; APPLICANT: Zhou, Yihua			
; APPLICANT: Cao, Yongwei			
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With			
; TITLE OF INVENTION: Plants			
; FILE REFERENCE: 38-21(53222)B			
; CURRENT APPLICATION NUMBER: US/10/425,115			
; CURRENT FILING DATE: 2003-04-28			
; NUMBER OF SEQ ID NOS: 369326			
; SEQ ID NO 139789			
; LENGTH: 2830			
; TYPE: DNA			
; ORGANISM: Zea mays			
; FEATURE:			
; OTHER INFORMATION: Clone ID: MRT4577_58974C.1			
US-10-425-115-139789			
Query Match 72.5%; Score 2495.8; DB 18; Length 2830;			
Best Local Similarity 95.1%; Pred. No. 0;			
Matches 2653; Conservative 0; Mismatches 102; Indels 36; Gaps 6;			
Qy	637	CGAGGTCTCGTCTTCCCTGCAACAGGGATCCATCCGTACCCCTGTGTCTGAGCCAGGGAG	696
Db	1	CGAGCTCTCGTCTTCCCTGCAACAGGGATCCATCCCTACCCCTGTGTCTGAGCCAGGGAG	60
Qy	697	TGCCAAGTGGGACGAGAAAGAAAGTAGTGAGCTGGGAAGGAGAGGATGGACGACTGGAAGTC	756
Db	61	CGCCAAAGTGGGACGAGAAAG-AGTGAGCTGGGAAGGAGAGATGGACGACTGGAAGTC	119
Qy	757	CAAGCAGGGCATCTCTCGCGGGCGC-----GCCGATCCCGAAGACATGGACGCCGAGCT	810
Db	120	CAAGCAGGGCATCTCTCGCGGGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	179
Qy	811	GGCACTGAACGACGAGGGGAGGACCGCTGTTCGGAAGAGGTGTGATCGCGTTCGAGCAA	870
Db	180	GCCGCTGAACGACGAGGGGAGGACCGCTGTTCGCGCAAGAGGTGTGATCGCGTTCGAGCAA	239
Qy	871	GGTGAACCCGTAACCGGATGGTGTATCGTGTGGGTCTCGTTGTGCTCGCCTCTTCTCTCG	930
Db	240	GGTGAACCCGTAACCGGATGGTGTATCGTGTGGGTCTCGTGTGCTCGCCTCTTCTCTCG	299
Qy	931	GTACCGTATCTGTGACCCCGTCCCGGACGCCATCCGGGCTGTGGCTGTCTCATCATCTG	990
Db	300	CTACCGCATCTGTGACCCCTGTCCCGGAGCCATCCGGGCTGTGGCTGTCTCATCATCTG	359

Qy	991	CGAGATCTGGTTCCGCATCTCTCTGGATCTCTCGACCAAGTTCCTCCCAAGTGGTTCCCAATCGA	1051
Db	360	CGAGATCTGGTTCCGCCGTGCTCTGGATCTCTCGACCAAGTTCCTCCCAAGTGGTTCCCAATCGA	419
Qy	1051	CCGCGAGACGTACTCGACCGCTCTCTCCCTCAGGTACGAGAGGAGAGGAGCGCTCGCT	1110
Db	420	CCGCGAGACGTACTCGACCGCTCTCCCTCAGGTACGAGAGGAGAGGAGCGCTCGCT	479
Qy	1111	GCTGTCCGCGGTGGACCTGTTCTGTGAGCAGCGGTGGACCCGCTCAAGAGAGCGCGCTGGT	1170
Db	480	GCTGTCCGCGGTGGACCTGTTCTGTGAGCAGCGGTGGACCCGCTCAAGAGAGCGCGCTGGT	539
Qy	1171	GACCGCAACAACCGTGTCTCTCCATCTCTGCCGTAGACTACCCCGTGGACAAGGTCTCTCTG	1230
Db	540	CACCGCAACAACCGTGTCTCTCCATCTCTGCCGTGGACTACCCCGTGGACAAGGTCTCTCTG	599
Qy	1231	CTACGTTCTCCGACGACGCGCGCTCGATGTGACGTTTCGAGTCTGCTCGAGAGCGCCGA	1290
Db	600	CTACGTTCTCCGACGACGCGCGCTCGATGTGACGTTTCGAGGCGCTGTCTGGAGACGCGCGA	659
Qy	1291	GTTTCGCGCGCAAGTGGGTGCCCTTCTGCAAGAAGTTCGGCATCAGAGCCCGCGCCCGCA	1350
Db	660	GTTTCGCGCGCAAGTGGGTGCCCTTCTGCAAGAAGTTCGGCATCAGAGCCCGCGCCCGCA	719
Qy	1351	GTTCTACTTCTCTCGTCAAGTTCGACTCTCAAGTTCGAGGTCAGCGCCACCTTCGTGCA	1410
Db	720	GTTCTACTTCTCTCGTCAAGTTCGATTACCTCAAGGACNAGGTTCAGGCCACCTTCGTGCA	779
Qy	1411	GGAGCGCGGCCATGNAAGAGAGATATGAGGAGTTCNAGGTTCGGATCCGATCAACCGGCTGGT	1470
Db	780	GGAGCGCGGCCCATGNAAGAGAGATATGAGGAGTTCNAGGTTCGGATCCGATCAACCGGCTGGT	839
Qy	1471	GGCCNAGGCCCATGNAAGTTCGCGCAGAGGGTTCGATCATGAAGGACGGCACGCGCTGGCC	1530
Db	840	GGCCNAGGCCCATGNAAGTTCGCGCAGAGGGTTCGATCATGAAGGACGGCACGCGCTGGCC	899
Qy	1531	CGGGAACAACACCCGCGACCAACCCCGGCATGATCCAGGTGTTCTCTGGGCCACACGCGCGG	1590
Db	900	CGGGAACAACACCCGCGACCAACCCCGGCATGATCCAGGTGTTCTCTGGGCCACACGCGCGG	959
Qy	1591	CCACGACACGAGGGGCAACGAGTTCGCCCGCTCTGTGTAGCTCTCCGCTGAGAAGCGCCC	1650
Db	960	CCACGACACGAGGGGCAACGAGTTCGCCCGCTCTGTGTAGCTCTCCGCTGAGAAGCGCCC	1019
Qy	1651	GGGATTCACGACCAACAAGAAGCCGCGGCATGAACGCTCTCATTTCCGCTCTCCGCGCT	1710
Db	1020	GGGCTTCCAGCAACAAGAAGCCGCGGCATGAACGCTCTCATTTCCGCTCTCCGCGCT	1079
Qy	1711	GCTGACCAACCGGCCATTATGTCTCAATTTGGACTGTGATCACTACATCAACAACAGCAA	1770
Db	1080	GCTGACCAACCGGCCCTTATGTCTCAACTTTGGACTGTGATCACTACATCAACAACAGCAA	1139
Qy	1771	GGCCATCCGGGAGGCCATGTGCTTCTCATGGACCCCTCAGGTTCGGCGGGAAGTCTGCTA	1830
Db	1140	GGCCATCCGGGAGGCCATGTGCTTCTCATGGACCCCTCAGGTTCGGCGGGAAGTCTGCTA	1199
Qy	1831	CGTTTCAGTTCGCCGAGAGGTTTCGACGCGCATCGACGCTGACACCGATACGCTAACAGGAA	1890
Db	1200	CGTGACGTTCCCTTCAGAGGTTTCGACGCGCATCGACGCTGACACCGATACGCGAACAGGAA	1259
Qy	1891	CACCGTCTTCTTCGACATCAACATGAAGGGGCTGGACGGCATCAAGAGCCCGGTGTACGT	1950
Db	1260	CACCGTCTTCTTCGACATCAACATGAAGGGGCTGGACGGCATCAAGAGCCCGGTGTACGT	1319
Qy	1951	CGGGAACAGGTGTGTTTCGGCGCCAGCGGCTCTACGGGCTACAACTCTCCAGGGAGCC	2010
Db	1320	CGGGAACAGGTGTGTTTCGGCGCCAGCGGCTCTACGGGCTACAACTCTCCAGGGAGCC	1379
Qy	2011	CAAGAGCCCAAGATGTTGACTTCGCACTCTCCCGTGTCTTCGGCGCGCAAGAGCGGAA	2070
Db	1380	CAAGAGCCCAAGATGTTGACTTCGCACTCTCTCCCGTGTCTTCGGCGCGCAAGAGCGGAA	1439
Qy	2071	ACACGCCAAGGAGCGGGCTGCGGAGGGGACCGCTGATATGGGAGTATAGCGACAAGGA	2130

Db 1440 GGACGCCAAGATGGCTGCCGAGGGGACCGCGGATATCGAGTGACAGCGCAAGGA 1499
Qy |||||
Db 2131 GATGCTCATGTCACACATGAATTCGAGAACGGTTTCGGGAGTCCCGGGGCTTCGTAC 2190
Qy |||||
Db 1500 GATGCTCATGTCACACATGAATTCGAGAACGGTTTCGGGAGTCCCGGGGCTTCGTAC 1559
Qy |||||
Db 2191 GTCACGCTGATGAGGAGGCGGGGTCCTCTTCGTTCGAGCCCGCGGGCTTCCTCAA 2250
Qy |||||
Db 1560 GTCACGCTGATGAGGAGGCGGGGTCCTCTTCGTTCGAGCCCGAGTTCGCTCTCAA 1619
Qy |||||
Db 2251 GGAGCCATCATGTCATCAGCTCGGCTACGAGCAAGACCGACTGGGGGCTGGAGCT 2310
Qy |||||
Db 1620 GGAGCCATCATGTCATCAGCTCGGCTACGAGCAAGACCGACTGGGGGCTGGAGCT 1679
Qy |||||
Db 2311 GGGGTGATCTACGGGTGATCAGCGAGGACATCTGACGGGGTTCAAGATGCACTGCCG 2370
Qy |||||
Db 1680 GGGGTGATCTACGGGTGATCAGCGAGGACATCTGACGGGGTTCAAGATGCACTGCCG 1739
Qy |||||
Db 2371 CGGTGGGCTCGGTGATGATGATGCGAGAGCGGGGCGGTTCAGGGGTCGGGCGCGAT 2430
Qy |||||
Db 1740 CGGTGGGCTCGGTGATGATGATGCGAGAGCGGGGCGGTTCAGGGGTCGGGCGCGAT 1799
Qy |||||
Db 2431 CAATCTATCGAGCGTCTCAACAGGTGCTCCGGTGGGCGTGGGTCCGTCGAGATCTT 2490
Qy |||||
Db 1800 CAACCTGTCCGACCGTCTCAACAGGTGCTCCGGTGGGCGTGGGTCCGTCGAGATCTT 1859
Qy |||||
Db 2491 CTTGAGCGGACAGCCCGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 2550
Qy |||||
Db 1860 CTTGAGCGGACAGCCCGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1919
Qy |||||
Db 2551 GCGCTTCGCTATCAACACCAATACCTTACCTTCGCTCGCTCGCTCGCTCGCTCGCT 2610
Qy |||||
Db 1920 GCGCTTCGCTATCAACACCAATACCTTACCTTCGCTCGCTCGCTCGCTCGCTCGCT 1979
Qy |||||
Db 2611 CTGACACCTCCCGCGTCTGCTCCCTCACCGGAGTTCAATCATGCGCTGATTCAGAC 2670
Qy |||||
Db 1980 CTGTACCTTCGACCGTCTGCTCCCTCACCGGAGTTCAATCATGCGCTGATTCAGAC 2039
Qy |||||
Db 2671 GTTCGCGACGCTCTTCTTCATCGCGCTCTTATGATGATGATGATGATGATGATGAT 2730
Qy |||||
Db 2040 GTTCGCGACGCTCTTCTTCATCGCGCTCTTATGATGATGATGATGATGATGATGAT 2099
Qy |||||
Db 2731 GATCGGTGAGCGGGGTGAGCATCGAGAGTGTGAGAGAACGAGCAGTTCTGGGTAT 2790
Qy |||||
Db 2100 GATCGGTGAGCGGGGTGAGCATCGAGAGTGTGAGAGAACGAGCAGTTCTGGGTAT 2159
Qy |||||
Db 2791 CGGCGGCTGTCGCGCATCTTTCGCGTGTGAGGGCTGCTCAAGTCTCTCGCGG 2850
Qy |||||
Db 2160 CGGCGGCTGTCGCGCATCTTTCGCGTGTGAGGGCTGCTCAAGTCTCTCGCGG 2219
Qy |||||
Db 2851 GATCGACACCAACTTTCACCGTCACTTCAAGGCCACCGGCGACGAGGACGAGTTTCG 2910
Qy |||||
Db 2220 GATCGACACCAACTTTCACCGTCACTTCAAGGCCACCGGCGACGAGGACGAGTTTCG 2279
Qy |||||
Db 2911 CGAGCTTACGCTTCAAGTGAACAGCTCTCTCATCCGCGCCACACAGCTGCTCATAT 2970
Qy |||||
Db 2280 CGAGCTTACGCTTCAAGTGAACAGCTCTCTCATCCGCGCCACACAGCTGCTCATAT 2339
Qy |||||
Db 2971 TACGTTATCGGCTGTCGGCGGATCTCGAGCCATCAACACCGGTACAGTCTCTG 3030
Qy |||||
Db 2340 CAACATCATCGGCGTGTGGCGGATCTTCGACGCCATCAACACCGGTACCAAGTCTG 2399
Qy |||||
Db 3031 GGGGCGCTCTTTCGGCAAGCTCTTTCGCGCTTCGCTCATCGTCACTCTACCGCTT 3090
Qy |||||
Db 2400 GGGGCGCTCTTTCGGCAAGCTCTTTCGCGCTTCGCTCATCGTCACTCTACCGCTT 2459
Qy |||||
Db 3091 CCTCAAGGGGCTATGAGGGGCGCAGAACAGGACGCCACCGTGTGTGTATCTGGTCTCAT 3150
Qy |||||
Db 2460 CCTCAAGGGGCTATGAGGGGCGCAGAACAGGACGCCACCGTGTGTGTATCTGGTCTCAT 2519
Qy |||||
Db 3151 TCTGCTGGCTCATCTTCTCCCTGCTCTGGGTGAGGATCGACCCCTTCATCTGTCAGGAC 3210
Qy |||||

Db 2520 GCTGCTGGCTCATCTTCTCCCTGCTCTGGGTGAGGATCGACCCCTTCATCTGTCAGGAC 2579
Qy |||||
Db 3211 CAAGGGCCCGAGCTCAGGAGTGTGGCATCAATTCGTGAGCTGTTTATTAAGGTTCAAA 3270
Qy |||||
Db 2580 CAAGGGCCCGAGCTCAGGAGTGTGGCATCAATTCGTGAGCTGTTTATTAAGGTTCAAA 2639
Qy |||||
Db 3271 A---TTCTGAGCTTGTGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3327
Qy |||||
Db 2640 AAGATTCTAGAGCTTGTGATAGGAG-AAAACAACAGTTTGAAGATTTTGTATTGTCTATG 2698
Qy |||||
Db 3328 GTGTC-----TGTAATGTTATGTATGTTACCCAGAAATGT-----CGGA 3362
Qy |||||
Db 2699 TTTTGGAGCTCATACGCTAGCTTTTGTGTAATGTTCTTGGACCCAGAAATGTTCGGCCGGA 2758
Qy |||||
Db 3363 CGAGGAATTCGAAACAAAGGACAAAGGTTTGATT 3393
Qy |||||
Db 2759 CGAGGAATTCGAAACAAAGGACAAAGGTTTGAAT 2789
Qy |||||

RESULT 7

US-10-209-059-25
; Sequence 25, Application US/10209059
; Publication No. US2003016383A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 3470
; TYPE: DNA
; ORGANISM: Zea mays
; US-10-209-059-25

Query Match 45.4%; Score 1563.4; DB 16; Length 3470;
Best Local Similarity 70.9%; Pred. No. 0;
Matches 2369; Conservative 0; Mismatches 671; Indels 303; Gaps 10;

Qy 97 GGCCAGCGCGCGCTGGTGGCCGGCTCGCAACCCGGAACGAGCTGTGTGATTCGGGG 156
Db 43 GGTCAACCGGTGGCTCGCCCGGCTCGCACATGCGGAGCGAGCTGATGTCATGCGCGC 102
Qy 157 CCACGAGGACCCCAAGCC---GCTGGGGCGCTGAGCGGGCGAGTGGCGAGATATCGG 213
Db 103 CCAGGAGGAGCGCAACCGCAAGGTCCGAGCGCGAGCTGGAAGACGTCGCGGTGCGC 162
Qy 214 CGACGAGGTGCGGTCAACGCTGAGCGCGACCTTTCGTGCGCTGCAACGAGTGCAGCTT 273
Db 163 CGACGAGGTGCGGACGCGGAGGAGCGGCGAGCTTTCGTGCGCTGCGCGAGTGGGCTT 222
Qy 274 CCCCGTGTGCGCCCTTGTCTACGAGTACGAGCGCGGGAGGCAACGAGAACTGCCCCCA 333
Db 223 CCCCGTGTGCGCCCTTGTCTACGAGTACGAGCGAGCGAGGAGCGAGTGTGCGCGCA 282
Qy 334 GTCAGACGCGCTACAGCGCTCAAGGAGCGCGGAGGTTGCCGGGACGATGACGA 393
Db 283 GTCAACACCCGCTACAGCGCGAGAAAGGTTGCCGAGGTTGGAAGGAGCGAGGAGGA 342
Qy 394 GG---AGGACATCGACGACCTGAGGACGAGTTTCAACATCGACGAGAAATCAGCAGAG 450
Db 343 GGGCCCGGAGATGAGACGACTTCGAGGACGAGTTTCCCGCCCAAGAGCCCTCA 402

QY 451 GCAGCTGGAGGCAACATGACAGAACAGCCAGCATCAACGAGGCGATGCTGCAACGCGAGGAT 510
DB 403 CGAGCCTGTGCGCTTCGACGCTTACTCGAGAAACGCGACACCCGCGCGAGAAATGGCG 462
QY 511 GAGCTACGGGAGGGCCCGACGACGCGGACGGCAACAAACCCCGCAGATCCCGCCCAT 570
DB 463 GACGGGTGG-----CCAGACGCTGTCGT 486
QY 571 CATCACCGGCTCCGCTCCGTCGCGGTGAGCGGTGAGTTCCGATTACCAACGGGTATGG 630
DB 487 CTTACCGGAAGCGTCGCGGGGAGG----- 512
QY 631 CCAGGGGAGGTCGTGCTTCCTCCCTGCAACAAGCGATCCATCCGTACCTGTGTGTAGCC 690
DB 513 -----ACCTGAGGCGGAGAG 528
QY 691 AGGAGGTGCCAAGTGGGACGAGAAGAGTGAAGTGAAGGAGGAGGATGACCACTG 750
DB 529 GGAGATGGAGGGAGCATGAGTGGAGGACCGGATCGACAAGTGAAGACCAAGCAGGA 588
QY 751 GAAAGTCCAAAGCAGGGCATCTCGCGCGCGCGCGCGATCCCGAAGACATGAGACGCCGACGT 810
DB 589 GAAGAGGGCAAGCTCAACCAACGACGACGACGACGACGACGACGACGACGACGACGACG 648
QY 811 G----GCACCTGAACGACGAGCGGAGGACGCGCTGTGAGGAAAGGTGTGATTCGCGTCGAG 867
DB 649 GTACATGCTGCTTGGCGAGGCGCGACAGCGCGTGTGGCGCAAGGTTCCGATCCCGTCGAG 708
QY 868 CRAAGGTGAACCCCTACCGGATGATGCGGTGCGTCTGTTGTGCTGCCCTTCTTCT 927
DB 709 CATGATCAACCCGTACCGCATGTCATGCTGCTCCGCTGGTGGTGTCTGCTTCTTCTCT 768
QY 928 CCGGTACCGTATCCTGCACCCCGTCCCGGACGCGCATCGGGCTGTGGCTCGTCTCCATCAT 987
DB 769 CAAGTTCGGATACAGACGCCCGCCACGAGCGCGTGCCTCTGTGGCTGGCTCCGTCTAT 828
QY 988 CTCGGAGATCTGTTGCGCATCTCTGATCTCTGATCTCGACCAAGTTCGCCAAGTGGTTCGCCAT 1047
DB 829 CTCGGAGCTGTGTTGCGCTTCTCTGATCTCTGATCTCTGACCAAGTGGCGCGCGGT 888
QY 1048 CGACCGGAGACCTACTCGACCGCTCTCCCTCAGGTACGAGAGGAGGGAGCGGTC 1107
DB 889 GACCGGGAGACGTACTTGACCGCTTGGCGCTGCGGTACGACCGTGAAGGCGAGCGGTG 948
QY 1108 GCTGCTGTGCGCGGTGACCTGTTCTGTGACGACGCTGACCCGCTCAAGGAGCGCGCGT 1167
DB 949 CCGGCTGTCCCATCGACTTCTTGTGACACGCTGGACCCGCTCAAGGAGCGGCCAT 1008
QY 1168 GGTGACCGCCAAACACGCTGCTCTCATCTCGCGGTAGACTACCCCGTGAACAAGGTCTC 1227
DB 1009 CATACCGCCAAACACCGTGTGTCTCATCTCTCGCGCTCGACTACCCCGTGAACCGCTCAG 1068
QY 1228 CTCGTAGCTCTCGACGACGCGCGTCCGATCGTTCGAGTTCGCTGTCGAGACGCGC 1287
DB 1069 CTCGTAGCTCTCGACGACGCGCGTCCATGCTGCTCTCGACGCGTGTTCGAGACCGC 1128
QY 1288 CGAGTTTCGCGCGCAAGTGGTGCCTTCTGCAAGAAAGTTCGGCATCGAGCCCGCGCCCGC 1347
DB 1129 CGAGTTTCGCGCGCGCTGGTGCCTTCTGCAAGAAAGTTCGGCATCGAGCCCGCGCGCC 1188
QY 1348 GGAAGTTCTATTCTCGCTCAAGTGCATCTACCTCAAGGACAAAGGTGAGCCCACTTCGT 1407
DB 1189 GGAGTTCTATTCTCGCAGAAAGTTCGACTACCTCAAGGACAAAGGTGAGCCGACGTTCTG 1248
QY 1408 GCAGGAGCGCGCGCATGAAGAGAGTATGAGGAGTTCAAGGTCCGATCAACCGGT 1467
DB 1249 CAAGGAGCGCGCGCATGAAGAGGAGTACGAGGAGTTCAAGGTCCGATCAACCGGT 1308
QY 1468 GTTGGCCAAAGCCATGAAGTGC CGGACAGAGGGGTGGATCATGAAGGACGGCACGCGCTG 1527
DB 1309 GTTGGCCAAAGCCAGAGAAAGCCCGAGGAGGGTGGTCAATCAGGACGGCACGCGTG 1368
QY 1528 GCCCGGGAACAACACCGCGACCAACCCCGGCATGATCCAGGTTGTTCTCGGGCCACAGCGG 1587

DB 1369 GCCCGGGAACAACACGCGGACCAACCCGGGTATGATCCAGGTCTACTCGCAACCCAGGG 1428
QY 1588 CGGCCACGACACCGAGGGCAACGAGCTGCCCGCTCGTGTAGCTTCCCGTGAGAAGCG 1647
DB 1429 CCGCTTGAAGCTGGAGGGCCAGAGCTGCCGCGCTCTGTCTAGCTTCCCGTGAGAAGCG 1488
QY 1648 CCGGGATTTCACGACCAACAAGAAGCGCGCCATGAACGGCTCTGATTGCGGTCTCCCG 1707
DB 1489 CCGCGGTACACCAACCAACAAGAGGCGGCGCATGAACGGCTGGTGGCGTCTCCCG 1548
QY 1708 CGTGTGACAAACGCGCCATTTCATGCTCAACTTGGATCTGTGATCTACTATCAACAACAG 1767
DB 1549 CGTGTCTACCAACGCGCCCTTCATCTCAACCTCGACTCGACCACTACGTCAACAACAG 1608
QY 1768 CAAGGCCATCCGCGAGGCGCATGTTCTCTCATGACCTCTAGGTCCGCGGAAGTCTG 1827
DB 1609 CAAGGCGCTGCCGAGGCGCATGTTCTCTCATGACCCGAGCTGGGAAGAGTCTG 1668
QY 1828 CTACGTTCAAGTTCGCGCAGAGGTTTCGACGCGCATTCGACGTGACGACCGGATACGCTAACAG 1887
DB 1669 CTACGTTCAAGTTCGCGCAGCGCTTCGATGGCATCGATCGCCACGACCGATACGCCAACCG 1728
QY 1888 GAACACCGTCTTCTTCGACATCAACATGAAGGGGCTGGAACGATCCAAAGGCGCGGTGA 1947
DB 1729 CAACGTCGTCTTCTTCGACATCAACATGAAGGGGCTGGAACGATCCAGGGCGCGGTGA 1788
QY 1948 CGTGCGGACAGGTCGTCTCCGCGCCAGCGCTCTACGGCTACAACTCCCAAGGG 2007
DB 1789 CGTGCGGACGGGTGCGTGTTCAACCGCCAGGCGCTGTACGGCTACGACCGCGCGGCC 1848
QY 2008 ACCCAAGAGCGCCAAAGATGTTGACC-----TGCGACTGCTGCCCGCTG 2049
DB 1849 CGAGAAGCGGCCCAAGATGACGTGCGACTGCTGCGCGTCTGTGCTGTCTGCTGCTG 1908
QY 2050 CTTGCGCGCGCAAGAAGCG----- 2067
DB 1909 CTTGCGCGCGCAAGCGCGCAAGCGCGCAAGAACAAAGAGGCGACGCGCGCGAGGA 1968
QY 2068 -----GAAACACGCCAAGACCG 2085
DB 1969 GCGCGCGCGGGCTGTCTCGGCTTCTACAGGAAGCGGAGCAAGAGGAACAAGCTCGCGG 2028
QY 2086 GCTGCGGAGGGCACCGCTGATATGGAGTAGATAGCAACGAGGAG----- 2131
DB 2029 CCGGTGCGTGGCGCGGACGCAAGAGGGCGCGGTGTACAAGAAGCACACAGCGCGCTT 2088
QY 2132 -----ATGCT 2136
DB 2089 CGAGCTGGAGGAGATCGAGAGGGGCTGGAGGGGTACGACGAGCTGGAGCGCTCTCGCT 2148
QY 2137 CATGTCCACATAGACTTCGAGAGCGGTTCGGGACAGTCCGCGGCTTCGTCAAGTTCGAC 2196
DB 2149 CATGTGCAAGAGAGCTTCGAGAGCGGTTCGGCCAGTCCGCCGTGTTATCGCTCCAC 2208
QY 2197 GCTGATGAGGAGGAGCGCGTCCCTC-----CTTCTGTCGAGCCCGCGCTCCTCAA 2250
DB 2209 GCTGCTGAGGAGCGCGGCTTCGCGAGGCGCGCGCGCGCGCGCTCATCAA 2268
QY 2251 GGAGGCCATCAATGTCATCAGCTGCGGTACGAGGACAAAGACGATGGGGGCTGAGCT 2310
DB 2269 GGAGGCCATCCAGCTCATCAGCTGCGGATACGAGGAGAAAGCCAGTGGGGCAAGGAGAT 2328
QY 2311 GGGGTGGATCTACGGGTTCATCAAGGAGCATCTCTGACGGGTTCAAGATGCACATGCCG 2370
DB 2329 TGGGTGGATCTATGGGTTCGATCAGAGGATATCTCTGACGGGTTCAAGATGCATGCCG 2388
QY 2371 CCGGTGGCGCTCCGTGATCTGATGCCGAAGGGCGCGGTTCAGAGGGTTCGCGCGCGAT 2430
DB 2389 GGGGTGAAGTCCGTGATCTGCAACCGCACCGCGCGGTTCAGAGGGTTCGCGCGCCAT 2448
QY 2431 CAATCTATCGGACCGTCTCAACAGGTGCTCCGTTGGGCGCTGGGGTCTCGTCCAGATCTT 2490

Db 2449 CAACCTTGCTGATCGTCTCCACAGGTGCTGGCGCTGGGCTGGGTCCGTGGAGATCTT 2508
QY 2491 CTTTACGGCGGACACGCCCTCTGCTGTACGGCTTACGAAGAACGGCAACCTCAAGTGGCTCGA 2550
Db 2509 CATGAGCGCCACTCGCGCTCCGTACGCCCTAC---GGCGCCCGGCTCAAGTGGCTCGA 2565
QY 2551 GCGCTTCGCTTACATCAACACCATCTACCCCTTACCTTCCCTCGCTCCGCTCGCCTA 2610
Db 2566 GCGCTTCGCTTACATCAACACCATCTGTGTACCCCTTACCTTCCCTCGCTCCGCTA 2625
QY 2611 CTGCAACCTCCCGCGCTGCTCTCTCACCGGCAAGTTTCATCATGCCGTGATAGCAC 2670
Db 2626 CTGCAACATCCCGCGCTGCTGCTCTCACCGGCAAGTTTCATCATGCCGTGATAGCAC 2685
QY 2671 GTTCCGCAAGCTCTTCTTCAATGCTGCTCTTCAATGCTTTCGCAAGGGGATCTCGA 2730
Db 2686 CCGCTTCGCTTACATCAACACCATCTGTGTACCCCTTACCTTCCCTCGCTCCGCTA 2745
QY 2731 GATCGGTGGAGCGGGGTGAGCATCGAGAGGTGGTGGAGGACGAGCTTCTGGGTCTAT 2790
Db 2746 GCTGGGTGGAGCGGGGTGAGCATCGAGACTGTGTGGCGCAACGAGAGCTTCTGGGTCTAT 2805
QY 2791 CGCGGGCTGTCCGCGCATCTTCCGCGTCTGTGAGGGCTCTCAAGGCTCTCGCGG 2850
Db 2806 CGCGGGCTGTCCGCGCATCTTCCGCGTGTTCAGGGCTTCTCAAGGTTCTGGGCG 2865
QY 2851 GATGACACCAACTTCAACGCTCACTCAAGGCCACCGCGCAGAGGACGAGTTCCG 2910
Db 2866 CGTGACACCAAGCTTCAACGCTCACTCAAGGCCCGCGGACGAGGCGGCGCTTCG 2925
QY 2911 CGAGCTTACGCTTCAAGTGACACAGCTCTCTATCCGCGCTCTCAAGGCTCTCATCAT 2970
Db 2926 GGAACCTTCACTTCAAGTGACACAGCTCTCTATCCGCGCTCTCAAGGCTCTCATCAT 2985
QY 2971 TAACTGTATCGCGCTGTGGCGGATCTCCGAGCCCATCAACACGGGTACAGTCTCTG 3030
Db 2986 CAACATGTGGGATCTGTGGCGGCTGTCCGAGCGCTCAACACGGGTACAGGCTCTG 3045
QY 3031 GGGGCGCTCTTCGCAAGCTCTTCTCGCTTCTGGGTCTATGCTTCTGCTTCTGCTTCT 3090
Db 3046 GGGGCGCTCTTCGCAAGCTCTTCTCTCTCTTCTGGGTCTATGCTTCTGCTTCTGCT 3105
QY 3091 CTTAAGGGCTCATGGGGCGGCGAGAACAGGACGCGGCTTGTGTCTATCTGCTTCT 3150
Db 3106 CTTAAGGGGCTCATGGGGAGGAGAGACCGGACGCGGCTTGTGTCTGCTTCTGCTTCT 3165
QY 3151 TCTGTGCTCTCATCTTCTCTCTCTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 3210
Db 3166 CTTCTCTGCTTCTCTCTCTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 3225
QY 3211 CAAGGGCGGAGCTCAGGAGTGTGGCATCAATGCTGAGCT 3253
Db 3226 CAAGGGCGGAGCTCAGGAGTGTGGCATCAATGCTGAGCT 3268

RESULT 8

US-10-627-132-25
; Sequence 25, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; TITLE OF INVENTION: Theof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 3470
; TYPE: DNA
; ORGANISM: Zea mays
US-10-627-132-25

Query Match 45.4%; Score 1563.4; DB 17; Length 3470;
Best Local Similarity 70.9%; Pred. No. 0;
Matches 2369; Conservative 0; Mismatches 671; Indels 303; Gaps 10;
QY 97 GGCAGAGCGCGGGCTGTGTGGCGGCTCGCACAAACCGGAAAGAGTGTGTGTGATCGGGG 156
Db 43 GGTCAACGGGTGCTCGCGCGGCTCGCACATCGGGGACGAGTGTGTGTGATCGGGG 102
QY 157 CCAGAGGACCCCAAGCC---GCTGCCGGCGCTGAGCGGGCAGGTGCGAGATATCGG 213
Db 103 CCGGAGGAGCGCGAAGCCCAAGGTCCCGAGCGCCGACGTGAAGACGTGCCCGGTGTGCG 162
QY 214 CGAGAGGTGGGCTCAAGGTGAGCGGCGACCTCTTCTGCGCTTGCACACGAGTGGGCTT 273
Db 163 CGAGAGGTGGGCTCAAGGTGAGCGGCGACGCGGCGGCTTCTGCGGTGCGCGGCTT 222
QY 274 CCGGCTGTGGCGGCTTCTGCTACGAGTACGAGCGCGGAGGCGACGAGAACTGCGCCCA 333
Db 223 CCGGCTGTGGCGGCTTCTGCTACGAGTACGAGCGCGGAGGCGACGAGTGTGTGCGGCA 282
QY 334 GTGCAAGACGGCTTACAGCGCTCAAGGGGCGCGGAGGTTGCGGGGACGATGAGGA 393
Db 283 GTGCAACACCGCTTACAGCGCGCGAAGGGTGCCTCGAGGGTGAAGGGGACGAGGAGGA 342
QY 394 GG---AGGACATCGACGACCTTGAGCGACGAGTTCAACATCGACACGAGATCAGCAGAG 450
Db 343 GGGCGCGGAGTGAAGGCTTTCGAGGAGGAGTTCCTCCCGCAAGAGCGCCCAAGAGCTCA 402
QY 451 GCAGCTGGAGGCAACATGACAGAAACGACAGATCACGAGGCGATGTGTGACGCGGAGAT 510
Db 403 CGAGCTGTGCGGTTTCGAGCTTACTCGGAGAACGGGAGCACCGCGCGCAGAAATGCGG 462
QY 511 GAGCTACGGGAGGGGCGCGGACGAGCGGCGGCAACACACCGCGGAGATCCGCGCAT 570
Db 463 GACGGGTGG-----CCAGACGCTGTCTGTC 486
QY 571 CATCACCGGCTTCCGCTCCGCTGCGGTGAGCGGTGAGTTTCCGATTACCAACGGGTATGG 630
Db 487 CTTTACCGGAGCGTCCGCGGAGG----- 512
QY 631 CCAGCGGAGGTCTCTGTTTCTTCCCTGCAAGCGCATCCATCCGTACCTGTGTGTGAGCC 690
Db 513 -----ACCTGGAGGGGAGAG 528
QY 691 AGGAGTGCACAGTGGGACGAGAAAGAGTGAAGTGAAGTGGAGGAGGATGGACGATG 750
Db 529 GGAGATGAGGGGAGCATGTGAGTGAAGGACCGGATCGACAAAGTGAAGACCAAGCAGGA 588
QY 751 GAAGTCCAAGCAGGGCATCTTCCGCGCGCGCGCATCCGAGACATCGACGCGCGCAT 810
Db 589 GAAGAGGGGCAAGCTCAACACGACGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 648
QY 811 G---GCACTGAACGACGAGGCGAGCGCGCTGTGAGGAGAGTGTGATGCGGTGCGAG 867
Db 649 GTACATGCTGCTTCCCGAGGCGCGACAGCGCTGTGCGCGCAAGGTTCCGATCCCGTCGAG 708
QY 868 CAAGGTGAACCGGTACCGGATGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTTCTT 927
Db 709 CATGATCAACCGGTACCGCATCGTATCGTGTCTCCGCTGCTGCTGCTGCTGCTTCTTCT 768
QY 928 CCGGTACCGTATCTCTGACACCGCTCCCGGACCGCATCGGGCTGTGGCTGTCTCCATCAT 987
Db 769 CAAGTTCGGATCAGACGCGCGCCACGAGCGCGCTGCTCTGTGGCTTGGGCTTGGTCTAT 828


```
Db 2986 CAACATGTGGGATCTGTGGCGGCTGTCCGACGCCCTCAACACGGCTACGGCTCCTG 3045
Qy 3031 GGGGCCCCCTCTTGGCAAGCTCTTCTTGGCCCTTCTGGGTCTATCGTCCACCTCTACCCGTT 3090
Db 3046 GGGCCCGCTCTTGGCAAGCTCTTCTTCTTCTTCTTGGGTCTATCGTCCACCTCTACCCGTT 3105
Qy 3091 CCTCAAGGGCTCATGGGGCGCCAGAAACAGACGCCACCGTTGTTGTTCATCTGGTCCAT 3150
Db 3106 CCTCAAGGGCTCATGGGGCGCCAGAAACCGACGCCACCACTCGTCTGGTCCAT 3165
Qy 3151 TCTGTGCGCCCTCCATCTTCTCCCTCTCTGGGTCAAGATCGACCCCTTTCATCTCGTCAGAC 3210
Db 3166 CCTCTCGCCCTCCATCTTCTCTGCTCTTGGGTCAAGATCGACCCCGTTTATCCCGAAGGC 3225
Qy 3211 CAAGGCCCCGACGTCAGGCAAGTGTGGCATCAATGTCTGAGCT 3253
Db 3226 CAAGGCCCCCATCTCTCAAGCCATCGGAGTCGAGTCTGAGCT 3268

RESULT 9
US-09-900-237-13
; Sequence 13, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Syntheses
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 3517
; TYPE: DNA
; ORGANISM: Glycine max
US-09-900-237-13

Query Match 44.2%; Score 1523.2; DB 9; Length 3517;
Best Local Similarity 69.4%; Pred. No. 0;
Matches 2196; Conservative 0; Mismatches 918; Indels 52; Gaps 7;

Qy 92 ATGGAGGCGACGGCGGCTGGTGGCGGCTCGCAACCGGAACGAGCTGGTGTATC 151
Db 158 ATGGAAGCCAGCGCTGGACTGGTGGCTGACATACCGCAATGAGCTAGTTGTCA 217
Qy 152 CGGGGCGACAGGAGCCCAAGCGCTGGGGCGCTGAGCGGGCAGGTGTGCGAGATGC 211
Db 218 CATGGCATGAGAGCCGAGGCTTTGAAGAACTTGGATGGGCAAGTGTGAGATTGT 277
Qy 212 GCGCAAGAGGTGCGGCTCAAGCGGACCGCACTCTTCTGCGCTGCAACAGAGTGGCG 271
Db 278 GGTGATGCGGTGGGACTCAAGCGGTGATGAGACTTTGTTGGCTTGGCAATGAGTGGT 337
Qy 272 TTCCCGGTGTCGGCCCTGTACAGTACAGCGCCGGGAGGGACGACGAACTGCCCC 331
Db 338 TTTCCAGTGTGAGGCGCTTGTATGAGTATGAAAGGAGAGAAAGCACTTTTGCCCA 397
Qy 332 CAGTGAAGAGCGCTACAAGCGCTCAAGCGGAGCCGAGGTTGCCGGGACGATGAC 391
Db 398 CAGTGAAGAGCGCTACAAGCGCTCAAGCGGAGCCGAGGTTGCCGGGAGATGATG 457
Qy 392 GAGGAGGACATCGAGCATCTGGAGCAGGATTCACATCGACGACGAGAAATCAGCAGG 451
Db 458 GAAGAGGATGTGGATGATTTGAGCATGAAATCAATATGA----- 498
Qy 452 CAGCTGGAGGGCAACATCGCAACAGCCAGATCAACCGAGGCGATGCTGCACGGCAGGATG 511
```

```
Db 499 -----TGACAAAGAAACAAGCATGGCCAGGTTGCAAGCCATGCTTCATGGGAGGATG 553
Qy 512 AGCTTACGGGAGGGCCCGCGACGCGCGAAGCAACCAACCCCGCAGATCCCGCCCATC 571
Db 554 AGCTATGGAAGAGGTCTGAAGATGATG-----ACAATTCACAGTTCGCCAACACCTGTC 607
Qy 572 ATCAACCGGCTCCCGCTCCGTCGCGGTGAGCGGTGAGTTCGGAT---TACCAACGGGTAT 628
Db 608 ATTGCTGTGTGCTCTTAGGCGCTGTAAAGTGGGAGTTCCCAATATCATCTAAATGCTTAT 667
Qy 629 GGGCACGCGGAGGTCTGCTCTCCCTGCAAGCGCATCCATCCGTACCCCTGCTGTCGAG 688
Db 668 GGGGATCAGATGTTATCTCTTCACTGATAAAGAGTGCATCATATCCAGTGTCTGAA 727
Qy 689 CCAAGGAGTGCCTAAGTGGGACG-AGAAAGAAAGTGAAGTGGAGGAGAGATGAGCGA 747
Db 728 CCTGGAAGTGAAGTGGGACGAAAGAAAGAGATGGATGGAAGATAGATAGATGATGA 787
Qy 748 CTGGAAGTCAAGCAGGGGCTCTCGGGGGGGCGCGCATCCCGAAGACATGAGACGCGA 807
Db 788 CTGGAATTTGCAAGGCAATTTTGGGGCTTGAACCGGAT-----GAAGATCCAGA 838
Qy 808 CGTGGCACTGAACGACGAGGCGGACGCGCTGTCGAGGAAGTGTGATCGCGTCGAG 867
Db 839 TGCAGCCATGTTAGATGAAGCAAGGCAACCACTGTCAAGGAAAGTGCATATGATCCAG 898
Qy 868 CAAGGTGAACCCCGTACCGGATGGTGAATGCTGGTGGCTCTGTTGCTCGCTTCTTCCCT 927
Db 899 CAAATCAATCCATATAGATGGTGAATTTGGCAGCGTCTGGTATTCTTGTCTTCTTCCCT 958
Qy 928 CCGGTACCGTATCTGCAACCCCGTCCGAGCGCATCGGGCTGFGGCTGCTCATCAT 987
Db 959 CAGATACAGACTCATGAACCCAGTACATGATGATGCGCTGGGGCTATGGCTTAACCTCATCAT 1018
Qy 988 CTGCGAGATCTGGTTCGCCATCTCTGGATCTCTGACCATCTCCCAAGTGGTTCGCCAT 1047
Db 1019 ATGTGAATCTGGTTCCTTTTCAAGATCTCGATCAGTTTCCCAAAATGGTGTTCCTCAT 1078
Qy 1048 CGACCGCGAGACGTACCTCGACCGCTCTCCCTCAGGTACGAGGAGGAGGAGCGGCTC 1107
Db 1079 TGATAGAGAGACCTACCTTGCACCGCTCTTCCATCAGGTATGAGCGTGAAGTGAACCAA 1138
Qy 1108 GCTGCTCTCGGGGTGACCTGTTGTGAGACGCGTGGACCGCTCAAGAGAGCGCGCT 1167
Db 1139 CATGCTGCTCTGTAGATGTTTGTGTAGTACCGTGGATCCCATGAAGAACTCTCTCT 1198
Qy 1168 GGTGACCGCAACACCGTGTCTCCATCTCTCGGTAGACTACCCCGTGGACAGGTCTC 1227
Db 1199 GGTACAGCAACACTGTTCTTTCAATCTTTGGCCATGGATTTACCCGTTGATAAATATC 1258
Qy 1228 CTGCTAGCTCTCGACGACGCGCGCTCGATGCTGACGTTTCAAGTTCGCTGCGAGAGCGGC 1287
Db 1259 ATGCTACATTTCTGATGATGAGGCGCTCAATGTGTACATTTGAGTCTTATCAGAACTGC 1318
Qy 1288 CGAGTTCGCGCAAGTGGGTGCGCTTCTGCAAGAAAGTTCGGCATCGAGCCCGCGGCC 1347
Db 1319 AGAGTTTGTCTAGAAAGTGGGTACCGTTTGTGAAGAAATTTTCCATAGAACCTCGGGCAC 1378
Qy 1348 GAGTTCCTACTTCTCGCTCAAGGTGCGTACCTCAGGACGAGTTCAGGCGCGACCTTCGT 1407
Db 1379 TGAGATGATCTTCAGCGAGAGATGTTGACTACCTTAAAGGACAAAGTGCACACCTTTGT 1438
Qy 1408 GCAGGAGCGCGCGCCATGAAGAGAGATGATGAGGAGTTCAAGTTCGGATCAACCGGCT 1467
Db 1439 TAAGGAGCTGAGCTATGAAGGAGGATACGAGAGTTTAAGTTAGGATCAATGCACT 1498
Qy 1468 GGTGGCAAGCCATGAAGGTGCGGAGAGGAGTGGATCATGAAGACGCGACCGCTG 1527
Db 1499 TGTGCTTAAGGCCCGAGAAAGTTCTCAGGGAGGATGGATCATGAGGATGGGACACCATG 1558
Qy 1528 GCGCGGAGCAACACCGCGGACCAACCCCGGCGATGATCAGGTGTTCTCTGGGCCACAGGG 1587
```

Db 1559 GCCAGGGAATAACATAAGGATCATCTGGTATGATTAACAAGTGTTCCTTGGTAGCAGTGG 1618
Qy 1588 CGGCCACGACACCGAGGCAACGAGCTGCCCGCTCGTGTAGCTCTCCCGTGGAGAAGCG 1647
Db 1619 AGGTCTTGATCTGAAGGAACCAACTTCTCGCTTGTTATGTTTCCAGAGAGAAAG 1678
Qy 1648 CCGGGGATTCAGACCAACAGAAGGCGCGGCATGAACCGCTCTGATTCGCTGCCG 1707
Db 1679 GCCTGGTGTTCACACCAAGAAGCTGGTGCCATGAATGCTCTGCTCGGGTATCTGC 1738
Qy 1708 CGTCTGACCAACGCGCATTCATGCTCAACTTGGACTGTGATCACTACATCAACAACAG 1767
Db 1739 TGTCTCACAAATGCTCTTTCATGTTGAATCTGGATTTGATCACTATATCAATAACAG 1798
Qy 1768 CAAGGCCATCCGGGAGGCCATGTCTTCTCATGGAACCTCAGGTCCGCGGAAGGTCTG 1827
Db 1799 CAAGGCTGCCGAGAGGCCATGTCTTCTTGATGGACCCCAAACTGGGAAGAGGTCTG 1858
Qy 1828 CTAGTTCAGTTCCTCCGAGAGGTTGACGGCATCGAGCTGCAAGCCGATACGCTAACAG 1887
Db 1859 CTATGTCAGTTCCTCAAGAATTTGATGATGATGATGATACACATGATCGTTATGCCAACAG 1918
Qy 1888 GAACACGCTCTTCTTCACATCAACATGAAGGGCTGGAGGCATCCAAGGCCGGGTGA 1947
Db 1919 GAACACAGTTCCTTCTGATATTAACATGAAGGGCTAGATGGTATTAAGGGTCTGTATA 1978
Qy 1948 CGTCGGGACAGGGTGCCTGTTCCGGCGCCAGGCGCTCTACGGCTACAAACCTCCCAAGGG 2007
Db 1979 TGTGGGACTGGATGTGTTTCAGGAGCGAAGCTTTGTATGGCTATATCTCTCCAAAGG 2038
Qy 2008 ACCAAGAGGCCCAAGATGTGACCTCGACTGCTGCCCGTGTTCG ---GCCGCAAGAA 2064
Db 2039 TCCAAAGCGTCCAAATAATGTTAGCTGTGATTTGTTGCCCGTGTTCGGAAGCCGCAAGAA 2098
Qy 2065 GCGGAACAGCCCAAGAGCGGGTCCGAGGGCCACCGCTGATATGGAGTAGATAGCGA 2124
Db 2099 GTATAAGGAAGAATGATGCAATGAGAGGGCTGCAAGCCCTAAAGGGATGATGATGA 2158
Qy 2125 CAAGGAGATGCTCATGTCCACATGAATTCGAGAAGCGGTTCCGGCAGTCCCGCGGCTT 2184
Db 2159 CAAAGAGGTGTGATGTCCCAATGAATTTGAGAAGAAATTTGGACAATCCTCTATTTT 2218
Qy 2185 CGTCACGTGACGCTGATGAGGAGAGCGCGCTCCCTCTTCTGTCGAGCCCGCGCGCT 2244
Db 2219 TGTGACTTCTACCTTGATGGAAGAGGGTGTGTGCTCTCTTCTCAAGTCCAGCTGCCCT 2278
Qy 2245 CCTCAAGAGGCCATCCATGTATCATCAGCTGGGCTACGAGCAAGACCCAGCTGGGGCT 2304
Db 2279 GCTTAAAGAGCCCAATCATGTGATTAGCTGTGGATATGAAGATAAAACTGAATGGGGACT 2338
Qy 2305 GGAGCTGGGTGGATCTACGGGTGATCAGGAGGACATCTGACGGGGTTCAAGATGCA 2364
Db 2339 TGAGCTGGTTGGATCTATGATCTATACAGAGATATTTCAACAGTGTTTAAGATGCA 2398
Qy 2365 CTGCGCGGGTGGCGCTCCGTGTAATGCAATGCCGAAGCGCGCGCTTCAAGGGGTCCG 2424
Db 2399 TTGCGGTGGGTGGAGTCCATTTATTTGATGTCGAAGAGAGCTGCATTCGAAGGGTACTGC 2458
Qy 2425 GCGATCAATCTATCGAGCGCTCTCAACAGGTGCTCCGGTGGCGCTGGGTCCGTCGA 2484
Db 2459 TCCTATCAACTGTTCAGATCGTCTCAACCAAGGTTCTTCGTGGGCATCTGGTTCCTATGA 2518
Qy 2485 GATCTTCTTCAGCGCGGACAGCCCTGCTGTACGGCTACAAGAAACGGCAACCTCAAGTG 2544
Db 2519 GATTTTCTTAGTCACCATTTGCCCTCTATGGTATGGCTTCAAGGAAGAAGCTAAGTG 2578
Qy 2545 GCTGGAGCGTTGCGCTTACATCAACACCAATACCCCTTCACTCGCTCCCGTGTCT 2604
Db 2579 GCTTGAGAGATTTGCCCTATGCAAAACAACTGTCTATCCATTCACCTCCATTCCTCTAGT 2638
Qy 2605 CGCTACTGACCTCCCGCGCTGCTCTCCACCGGCAAGTTCAATCATGCGGTGAT 2664
Db 2639 TGCCTACTGTATCTTCCAGAGTGTGTTTACTCACTGACAAAATTCATCATGCGCCGAT 2698

Qy 2665 TAGCAGTTCGCGCAGCCTCTTCTTTCATCGCCCTCTTTCATGTCCATCTTCGCGACGGGCAT 2724
Db 2699 TAGCACCCTTGTCTGTTTGTACTTTGTCTCTCTCTCTCTCAATCATTTGCAACTGGTAT 2758
Qy 2725 CTTGGAGATCGCGGTGAGCGGGTGAAGCATCGAGGAGTGGTGGAGAAAGAGCAGTTCTG 2784
Db 2759 TCTTGAGTTCAAAATGGAGTGGAGCATTTAGGAAATGGTGAGAAATGAGCAGTTTTG 2818
Qy 2785 GGTTCATCGCGGGTGTCCGCGCATCTCTTCGCGTCGTCAGGGCTCTCTCAAGGTCTCT 2844
Db 2819 GGTTCATTTGGTGTGTATCAGCTCACCTCTTTGCTGTATACAGGTCTGCTAAAGGTCTCT 2878
Qy 2845 CGCGGGATCGACCAAACTTTCACCGTCACTCCCAAGGCCACCGCGGAGAGGACGACGA 2904
Db 2879 GGCTGGAATTGACACCAATTTCACTGTTACATCAAGGCAAC-----AGATGATGAAGA 2932
Qy 2905 GTTCGCGGAGCTCTACGCTTCAAGTGGAGCACGCTCTCTATCCCGCCCAACGAGTCTCT 2964
Db 2933 GTTTGGAGAAATTGTACACCTTTAAGTGGACTACACTCTTTGATTTCTTCCCAACCACTATTTT 2992
Qy 2965 CATCATTAACGTCATCGGCGTGTGGCGGCGCATCTCCGACGCCATCAACAAACGGGTACCA 3024
Db 2993 GATCATTAACATTTGTTGGTGTGTTGCTGGAATCTCAGATGCCATAAACAATGGGTACCA 3052
Qy 3025 GTCTCTGGGGCCCTCTTCGGCAAGCTCTTCTTCGCTTCTTCGGGTCTTCGTCACCTCTA 3084
Db 3053 ATCTCGGGGACCACTCTTTGGAAAGCTCTTCTTCTTCGTTGATTTGTCCTCTCTA 3112
Qy 3085 CCGGTTCTCAAGGGGCTCATGGGGGCCAGAAACAGGACGCCACCGTGTGTTGTCACTG 3144
Db 3113 TCCATTTCTTAAAGGTTTGATGGGTGCCAAATCGCAAAATCGCACACCCACCACTTTGTTG 3172
Qy 3145 GTTCATTTCTGGCTCCATCTTCTCCCTGCTCTGGGTGAGGATCGACCTTTTCATCGT 3204
Db 3173 GTCAGTGCTATTGGCTCTATTCTTCTTCTTCTTCTTGGTAAAGAAATGATTCATTTGCTCT 3232
Qy 3205 CAGGACCAAGGGGCCGAGCGTTCAGGCAAGTGTGGCATCAATTTGCTGA 3250
Db 3233 CAAGACTAAGGACCTGATACCAAGCTATGTGGAATCAACTGCTAA 3278

RESULT 10
US-10-437-963-83811/c
; Sequence 83811, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21 (53221) B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 83811
; LENGTH: 3493
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_83107C.1
US-10-437-963-83811

Query Match 41.8%; Score 1439.6; DB 18; Length 3493;
Best Local Similarity 69.4%; Pred. No. 0;
Matches 2296; Conservative 0; Mismatches 669; Indels 345; Gaps 11;

QY 187 GAGCGGCGAGTGTGCGGATATCGGCGAGAGGTGGGCTCAGGTTGGAGCGGCACCT 246
Db 3455 GAAGGAGAAGACGTGCCGGGTGTGCGGCGAGAGGTGGCGGAGGAGGCGGAAAGCC 3396
QY 247 CTTGTCGCTCCGACAGAGTGGCTTCCCGTGTGCGGCGCTGCTACAGAGTACGAGCG 306
Db 3395 GTTCGTGCGTGTGCGGAGTGGGCTTCCCGTGTGCGGCGCTGCTACAGAGTACGAGCG 3336
QY 307 CCGGAGGGGCGAGGAACTGCGCCCAAGTGAAGACGCGCTACAAAGCGCTTCAAAGGGAG 366
Db 3335 CAGCGAGGGGCAACCCAGTGTGCGCCCAAGTGAAGACGCGCTTCAAAGCGCCCAAAGGGGTG 3276
QY 367 CCCGAGGGTTCGCGGAGCGATGAGCA---GAGGACATCAGACCTGGAGCGAGGTT 423
Db 3275 CCCAGGGTGAAGGCGAGCGAGGACGACGGCGGCGACATGACGACTTCGAGGGAGGTT 3216
QY 424 CAACATCAGCAGCAGAGAAATGACGAGAGCGAGCTGGAGGGCAACATGCAAGAACAGCCAGAT 483
Db 3215 CCAGATCAAGAGCCCAACCAAGCAGA-----AACCCCCCAGAGCCCGTC 3170
QY 484 CACGAGGCGATGTGTGCAAGGAGATGAGCTACGGGAGGGGCCCCGACGACGCGGAGCG 543
Db 3169 AACTTCGACGCTACTCTCGGAGAAAGCGGAGCAGCGCGCACAGAAAGTGGCGCCCTGGAGG- 3111
QY 544 CAACAAACACCCCGCAGATCCCGCCCATCATCACCGGCTCCCGCTCCGTCGCGGTGAGCGG 603
Db 3110 -----CCCGGCGCTCTCTTCTTCAACCGGAAAGCGGTGGTGGGAGG- 3070
QY 604 TGAGTTCCGATTAACCAAGGATGAGCCACGGCGAGGTCTCGTCTTCCCTGCACAGCG 663
Db 3069 ----- 3070
QY 664 CATCCATCCGTAACCTGTGTGACCCAGGAGTGCCTAAGTGGGACGAGAAAGAAAGT 723
Db 3069 -----ATCTGAGCAGGAGAGGAGATGGAGGGTGGCAT---GGAGTGGAAAGGACAGAT 3018
QY 724 GAGCTGGAAGAGAGGATGAGCAGTGAAGTCCNAGCAGGGCATCTCCGCGGCGGCGC 783
Db 3017 CGAACAGTGGAAAGCAAGAGCAGGAGAAAGCGGGGCAAGTCAACCGCGACGACAGCGAGA 2958
QY 784 CGATCCCAAGACATGACGCGGACGTGGCACTGAACGACGAGGCGAGGACCGCTGTC 843
Db 2957 CGACGACGAAGAAGACGACGAGTACATGCTGCTCGGAGGCGAGGACGCGCTGTG 2898
QY 844 GAGAAAGTGTGATCGCGTCCAGCAAGTGAACCCGTAACCGGATGGTGTGATCGTGTGCG 903
Db 2897 GAGAAAGTGTGCGATCCCGTCCGAGCAAGATCAACCCGTAACCGGATCGTGTGCTCCG 2838
QY 904 TCTGTTGTGCTCGGCTTCTTCTCCGTAACCGTATCCTGACACCGCTCCGAGCGCAT 963
Db 2837 GCTGGTGTGCTGCTGCTTCTTCTCAAGTTCGCGATCACGACGCGGCGATGAGACGCGGT 2778
QY 964 CGGCTGTGGCTCGTCTCCATCATCTGCGAGATCTGTTTCGCCATCTCTGGATCCTCGA 1023
Db 2777 GCCGCTGTGGCTCGGTGATCTGCGAGCTGTGTTTCGCTGTGCTGTGATCTCTCGA 2718
QY 1024 CCAGTTCCTCAAGTGTGTTCCCATCGACCGGAGACGTACCTCGACCGGCTCTCCCTCAG 1083
Db 2717 CCAGTCCCAAGTGTGCGCGGTGACGAGGAGAGGTACCTGAGACGCGCTGGCCCTCCG 2658
QY 1084 GTACGAGGGAAGGGGAGCGTGTGCTGTGCGCGGTGAGACCTGTTCTGTGAGACGCT 1143
Db 2657 GTACGAGCGCAGCGGCGAGCGGTGCGCTTGGCCCCGATCGATTTCTTCGTGAGACGCT 2598
QY 1144 GGACCCGCTCAAGGAGCGCGCTGTGACCGCCAAACCGTGTCTTCCATCCTCGCGGT 1203
Db 2597 GGACCCGCTCAAGGAGCGCGCCATCATCACCGCAACACCGTGTCTGCTCATCTCGCGGT 2538
QY 1204 AGACTACCCCTGGACAAAGTCTCCTGTGATGCTTCCGACGACGCGGCTGATGCTGAC 1263
Db 2537 CGACTACCCCTGACCGCGCTCTCTGTGATGCTTCCGACGACGCGGCTCATGCTGCT 2478
QY 1264 GTTCGAGTGTGTCGGAGACGCGCGAGTTCCGCGCAAGTGGGGTGCCCTTCTGGAAGAA 1323

Db 2477 CTTGCAACGCTCTCCGAGACCGCCGAGTTGCGCCGCGGTGGGTCCCTTCTGCAAGAA 2418
QY 1324 GTTCGGCATCGAGCCCGCGCCCGGAGTTCTACTTCTCGCTCAAGGTTCGACTACCTCAA 1383
Db 2417 GTTCACCATCGAGCCCGCGCCCGGAGTTCTACTTCTCCAGAAGATCGACTACCTCAA 2358
QY 1384 GGAACAAGTGCAGCCCACTTCTGTGAGAGAGCGCGGCCCATGAAGAGAGATGAGGA 1443
Db 2357 GGAACAAGTGCAGCCCACTTCTGTCAAGAAAGCGCGGCCCATGAAGAGAGATGAGGA 2298
QY 1444 GTTCAAGGTCCGATCAACCGCTGTGGCCAAAGGCAATGAAGTGCAGAGGGGTG 1503
Db 2297 GTTCAAGGTGAGGATAAACCGCTGTGGCGAAGCGGAGAAAGCCGAGAGAGGGTG 2238
QY 1504 GATCATGAAGGAGCGCAGCGCTGTGGCCCGGAAACAACCCCGACCAACCCCGCATGAT 1563
Db 2237 GGTGATGACAGGAGCGCATGCGCGGGAAACAACAGGAGACCAACCCGGGATGAT 2178
QY 1564 CCAGGTGTTCTTGGGCCACAGCGCGGCCAGCACCGAGGGCAACAGAGTGCCTCCGCT 1623
Db 2177 CCAGGTGATCCTTGGGACAGCGAGCGCGCTCGACGTGAGGGCAGCGAGTGCCTCATG 2118
QY 1624 CGTGTAGTCTCCCGTGAAGAGCGCCGCGGATTCACGACACCAAGAGGCGCGGCCAT 1683
Db 2117 GGTGTAGTGTCCCGGAGAGGCGCCCGCTACAACCAAGAGGCGCGGCCAT 2058
QY 1684 GAACTCTCATCTCGCTCTCCGCTGTGCTGACCAAGCGGCAATCATGCTCAACTTGA 1743
Db 2057 GAATCCTCTGTTGCGCTCTCCGCGTGTACTACAACGCCCTTCTCATCTCAACCTGA 1998
QY 1744 CTGTGATCATATCATCAACAACAGAGGCAATCCGGAGGCCATGCTTCTTCTCATGGA 1803
Db 1997 CTGCGACCATCTAGCTCAACAACAGAGGCGCTCCGCGAGGCCATGCTTCTTCTCATGGA 1938
QY 1804 CCCTCAGTCCGCGGAGGTCTGCTAGTTCAGTTCGCGGAGGTTCCAGCGCATCGA 1863
Db 1937 CAAGCACTCGGCAAGAGCTGTGCTACGTCCAGTTCCTCCGCGCTTCGACGCGCATCGA 1878
QY 1864 CGTGCACGACCGATACGCTTAACAGAAACACGCTTCTTCTTC----- 1903
Db 1877 CCGCCAGATGCTGCTACGCCAACCGCAACCGTCTTCTTCGACGTGAGCACTGCTTCTT 1818
QY 1904 -----GACATCAACATGAAGGGCTGACGCGCATCCAAAG 1938
Db 1817 CTCGCGGTGCTGCTGCTGCGCGCTGACATCAACATGAAGGGCTGACGCGGATACAGG 1758
QY 1939 CCGGCTGTAGTTCGGGACAGGTGCTGCTGCTCCGCGGCGAGCGCTCTACGCTACACCC 1998
Db 1757 GCCGTTGTGCTGGGAGCGGGGACGCTGTTCAACAGGACGCGCTGTACGATACGACCC 1698
QY 1999 TCCCAAGGAGCCCAAGAGGCCAAGATGCTGAC-----TGCGATG 2040
Db 1697 GCCGCGCGGAGAGAGGCGGAGATGACGTGCGATGCTGCGCGCTGCTGCTGCTG 1638
QY 2041 CTGCGCTGTGCTGCGCGCAAGAGCG----- 2067
Db 1637 CTGCTGCTGCTTCCGCGGGGAGCGCGCAAGTTCGCAACAAGAGGGCGCGG 1578
QY 2068 ----- 2067
Db 1577 CCGCGCGAGGGGCGCGCTTCGACGAGCGCGCGCGGCTGCTCGGGTTCACAGAA 1518
QY 2068 -----GAAACAACGCAAGGACGGGCTCCGAGGGGACCGCTGATATATGGAGTAGATAGC 2122
Db 1517 GAGGAGCAAGAGAGCAAGCTCGCGCGGCGCGCTGCTGCTCGCGGAGGAGAAAGG 1458
QY 2123 GACAAGGA----- 2130
Db 1457 GTACCGGAAGCACGAGCGCGGTTCGAGCTGGAGGAGATCGAGGAGGCGCTCGAGGGGTA 1398
QY 2131 -----GATGCTCATGTCCCATGAACTTCGAGAGCGGTTTCGGGCA 2172

Db 1397 CGACGAGCTGAGCGCTCGTCGCTCATGTGCGAGAGAGCTTCGAGAGCGGTTCGGCCA 1338
Qy 2173 GTCGCGCGGTTCGTCACTGCGAGCTGATGAGAGAGCGGCGTCCCTC-----CTTC 2226
Db 1337 GTCGCGGTGTTTCATCGCTCCACCTCGTCGAGGACGGCGCTCCCGAGGGCGCGC 1278
Qy 2227 GTCGAGCCCGCGCGCTCTCTCAAGGAGGCAATCATGTATCATAGCTGCGGCTACGAGGA 2286
Db 1277 CGCGACCCCGCGCGCTCTCAAGGAGGCAATCATAGCTGCGGCTACGAGGA 1218
Qy 2287 CAAGACGACTGCGGGCTGAGCTGGGCTGGATCTACGGTTCGATCAGGAGGACATCCT 2346
Db 1217 GAAGACGAGTGGGCGAAGAGATTGGGTGATCTACGGTTCGATCGGAGGACATCTT 1158
Qy 2347 GACGGGTTCAGATGCACTCGCGGCTGGCGTCTCGTGTACTGTGATGCCGAGCGGC 2406
Db 1157 AACGGGTTCAGATGCAATTGGCGTGGTGGAGTTCGTGTACTGACGCGCGGAGGC 1098
Qy 2407 GGGTTCAGGGGTGCGCGCGGATCAATCTATCGGACCGTCTCAACAGGTGTCGCGTG 2466
Db 1097 GGCATTCAGGGGTGCGCGCGGCAATCAACCTGTGGATCGTCTGCACGAGTGTCCGGTG 1038
Qy 2467 GGGCTGGGTCCGTTCAGATCTCTTCAGCGGACAGCGGCTGTGTACGGCTACAA 2526
Db 1037 GGGCTGGGTCCGTTCAGATCTCTTCAGCGCGCAATTCAGCGCGCAATTCGGTACGCT---A 981
Qy 2527 GAACGGCAACCTCAAGTGGGTGAGCGCTTCGCTTACATCAACACCACTCTACCCCTT 2586
Db 980 TGGCGCGGCTCAAGTGGGTGAGCGCTTCGCTTACCAACCACTCTACCCCTT 921
Qy 2587 CACTCTCGCTCCGCTGCTCGCTACTGCACTCCCGCGCTGCTCTCTCAACCGGCAA 2646
Db 920 CACTCTCAATTCCTCGCTACTGCACTCCCGCGCTGCTCTCTCAACCGGCAA 861
Qy 2647 GTTCATCATCGCTCGATAGCAGTTCGCGAGCTCTTTCATCGCCCTCTTCATGTC 2706
Db 860 GTTCATCATCCCAACGCTTAACTTTGGCGAGATATGTTTATAGCGCTTTTCCTGTC 801
Qy 2707 CATCTTCGACCGGCACTCTCGAGATGCGGTGAGCGGCTGAGCATCAGAGATGGTG 2766
Db 800 GATCATCGACGGGGTGTGGAGCTGCGTGGAGCGGGGTGAGCATCAGGACTGGTG 741
Qy 2767 GAGGAACGAGAGTTCGTGGTCAATCGCGGCTGTCGCGGATCTCTTCGCGCTGTCGA 2826
Db 740 GAGGAACGAGAGTTCGTGGTGAATCGCGGCTGTCGCGGCACTGTTTCGCGCTGTTCCA 681
Qy 2827 GGGCTCTCAAGTCTCCCGGATCGACACCACTTCACGCTCACTCCAGGCCAC 2886
Db 680 GGGCTCTCAAGTCTCCCGGCTGGAACCACTTCACGCTGACGTCCAGGGCGC 621
Qy 2887 CGGCGACGAGGACGACGAGTTCGCGGAGCTCTACGCTTCAAGTGGACCACTCTCTCAT 2946
Db 620 CGCGACGAGACGAGCGGTTCGGGAGCTCTACCTGTTCAAGTGGAGAGCTGTGGT 561
Qy 2947 CCGGCCCAACGCTGCTCAATTAACGTAATCGCGGTGTCGCGGATCTTCGAGCC 3006
Db 560 CGCGGCCACGACGCTGATCATCAACATGTTGGGATCGTCCGCGGCTGTCGAGCC 501
Qy 3007 CATCAACAGGGTACAGTCTCGGGGCGCTCTTCGGAAGCTCTTTCGCTTCGCTTG 3066
Db 500 CGTGAACAAAGGGTACGGGTGCGGGGCGCTGTTGGGAAGCTCTTTCCTTCGCTG 441
Qy 3067 GGTTCATCGTCCACTTACCGTTCCTCAAGGGGCTCATGGGCGCGCAGAACAGGAGCC 3126
Db 440 GGTTCATCTCACTTACCGTTCCTCAAGGGGCTCATGGGAGGAGAACCGGAGCC 381
Qy 3127 CACGTTGTTGTCATCTGGTCCATTTGCTGGCTTCCATCTTCTCCCTGCTTGGGTGAG 3186
Db 380 CACCATTTGCTGCTGGTCCATCTCTCGCTCCATCTTCTCCCTGCTTGGGTGAG 321
Qy 3187 GATCGACCTTTTCATCTGAGGACCAAGGCGCGGAGTCAAGGCTGAGGATCAATTCG 3246
Db 320 GATCGACCTTTTCATCTCCCAAGCCCAAGGGCGCGCTCTCAAGGCTGCGGGGTCTCGTG 261

Qy 3247 CTGAGCTGTT 3256
Db 260 CTGAGCTGCT 251
RESULT 11
US-10-160-719-29
; Sequence 29, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 3746
; TYPE: DNA
; ORGANISM: Zea mays
; NAME/KEY: CDS
; LOCATION: (321)... (3551)
US-10-160-719-29

Query Match 40.6%; Score 1397.6; DB 16; Length 3746;
Best Local Similarity 67.2%; Pred. No. 0;
Matches 2185; Conservative 0; Mismatches 924; Indels 143; Gaps 9;
Qy 133 GAACGAGCTGTGCTGATCGGGGCCACGAGGACCCAAAGCGCTCGGGCGCTGAGCGG 192
Db 308 GAGTCTGCTTCCATGGAGGGCGACGCGGACGCGGTGAAGTCGGGAGGCGCGTGGCGG 367
Qy 193 GCAGGTGTGAGATATGCGGCGACGAGTTCGGGCTCACGGTGGACGCGGACCTTTCGT 252
Db 368 ACAGGTGTGAGATCTGCGGCGACGCGTGGGCAACGCGGAGGGGACGCTTCGC 427
Qy 253 CGCTTCACAGTGGCGGTTCCTCCGTGTGCGGCGCTGTACGAGTACGAGCGCGGGA 312
Db 428 CGCTTCGACGCTGTCGGGTTTCGGTGTGCGGCGCTGTACGAGTACGAGCGCAAGGA 487
Qy 313 GGGCACGAGACTGCCCCAGTGAAGACGCGCTACAAGCGCTCAAGGGGAGCCGAG 372
Db 488 CGGACGCGAGGCGTCCCCAGTGAAGACCAAGTGAAGCGCCACAGGGGAGCCGCGC 547
Qy 373 GGTTCGCGGGGACGATGACGAGGAGGACATCGACGACCTGGAGCACGAGTTCACATCGA 432
Db 548 GATCCGTGGGAGGAGGAGACGACATGATGCGGATAGCGACTTCAATTACCTTGCATC 607
Qy 433 CGACGAGATCAGCAGAGGCGTGGAGGGAACATGCGAGAACAGCCAGATCA----- 485
Db 608 TGGCAATGAGGACCAAGACAGAGATTGCCGACAGATGCGCAGATGCGCGCATGAACGT 667
Qy 486 -----CCGAGGCGATGCTGCGAGGAGTACGCGAGGGGCGCCCG----- 530
Db 668 TGGGGGACGCGGGATGTTGGTCCGCCAAGTATGACGTGGCGAATCGGGCTTACCAA 727
Qy 531 --ACGCGCGACGCGCAACAAACACCCCGCAGATCCCGCCCATCATCCCGGCTCCCGCTC 588
Db 728 GTATGACAGTGGCGAGATTCTCGGGGATATCTCCCATTCAGTCACTAAACAGCCAGATCTC 787
Qy 589 CGTCCGCTGAGCGGTGAGTTTCC--GATTACCAACCGGTTATGGCCACGCGGAGGTCTCGT 647

Db	788	AGGAGAAATCCCTGGTCTCCCTGACATCATATGATGTCCTCCCACTGGGAAATGG	847
Qy	648	CTTCCTCCACAAAGCGCATCCATCCGTATCCCTGTGTCTGAGCAGGAGTGCCAAAGTGGG	707
Db	848	CAAGCGTGCTCCATTTTC-CTATGTGAACCAATTCGCCAAATCCGTCAAGGGAGTTCTCTG	906
Qy	708	ACGAGAAAGAAAGTGTAGCTGGAAGGAGAGATGAGACACTGGAAGTCCAAAGCGGCA	767
Db	907	GTAGCATTTGGGAATGTTCCTGGAAGAGAGGGTTGATGGCTGGAAAAATGAACGACGACA	966
Qy	768	-----TCC	770
Db	967	AGGGACGATTTCCCATGACGAATGSCACAAGCAATGTCTCCCTCTGAGGGTGGGTGTG	1026
Qy	771	TCGGCGGGCGCCGATCCCGAAGACATGAGCGCGACGTGGCACTGAACGACGAGGCGA	830
Db	1027	GTGATATTGATGSCATCAACTGATTACAAATGGAAGATGCTTTATTTGAACGACGAAATC	1086
Qy	831	GGCAGCCGTGTCGAGGAAGGTGTCGATCGGTGAGCAAGGTGAACCCGTACCGGATGG	890
Db	1087	GACAGCCTTATCTAGGAAAGTTCCACTTCTCTCCAGGATAAATCCATACAGGATGG	1146
Qy	891	TGATCGTGTGGTCTCGTTGTGCTCGCTTCTCTCCGTGACCGTATCTGCAACCCCG	950
Db	1147	TCATTGTGCTGGATTGATTGTTCTAAGCACTTCTTGCACTACCGTATCAAAATCCTG	1206
Qy	951	TCCCGACGCCATCGGGCTGTGGTCTGCTCATCATCTGCGAGATCTGGTTCGCCATCT	1010
Db	1207	TGCGCAATGCAATACCAATTATGGCTTCTATCTGTTATATGTGAGATCTGGTTGCTCTT	1266
Qy	1011	CCTGGATCCTCGACAGTTCCCAAGTGTTCCTCCATCGACCGGACGATACCTCGACC	1070
Db	1267	CGTGGATATTGATCAGTTCCCTAAGTGGTTTCCAAATCAACCGGGAGAGCTACCTTGATA	1326
Qy	1071	GCCTCTCCTCAGGTACGAGAGGGAAGGGAGCGCTGCTGCTGTCGCGGTGGAACCTGT	1130
Db	1327	GGCTGGCAATTAAAGGTATGACCGGGAAGGTGAGCCATCTCAGTTGGCTGCTGTTGACATT	1386
Qy	1131	TCGTGAGACGGTGACCCGCTCAAGGAGCGCGCTGCTGAGACCGCCAAACCCGTGCTCT	1190
Db	1387	TCGTGAGTACAGTGCACCCATGAAGGAGCCTCTCTTGTCTCACTGCCAATACCGTGTAT	1446
Qy	1191	CAATCCTCGCGTAGACTACCCGTTGGAAGAGTCTCCTGCTAAGTCTCCGACGAGCGCG	1250
Db	1447	CAATTTCTGTGTGATTACCTGTGGATAAGGTCTCTTGTCTATGATATCTGATGATGAG	1506
Qy	1251	CGTCGATGTCAGCTTCAGTCTGCTGCGAGCGCGAGTTGCGCGCAAGTGGGTGC	1310
Db	1507	CTGCGATGCTGACATTTGATGCACTAGCTGAGACTTCAGAGTTTCTAGAAAAATGGGTAC	1566
Qy	1311	CCTTCTGCAAGAAGTTCCGCAATCGAGCCCGCGCCCGAGTTCTTACTTCTCGCTCAAGG	1370
Db	1567	CATTTGTTAAGAAGTACAACATTGAACCTAGAGCTCTCTGAATGGTACTTCTCCAGAAA	1626
Qy	1371	TCGACTACTTAAGACAAAGTGTGAGCCCACTTCGTGAGAGGCGCGCGCATGAAGA	1430
Db	1627	TTGATTACTTGAAGACAAAGTGCAACCTTCAATTTTAAAGACCGCGCGGCATGAAGA	1686
Qy	1431	GAGAGTATGAGAGTTCAAGGTTCGCGATCAACGCGCTGCTGCGCAAGGCCATGAAGTGC	1490
Db	1687	GAGATATGAAGNATTAAGTTAGGTTAAATGGCTTTGCTTAAAGCACAGAAATTC	1746
Qy	1491	CGGACAGGGGTGGATCATGAAGGACGGCAGCCGTGGCCCGGGAAACAAACCCCGACC	1550
Db	1747	CTGAGNAGGATGGATCATGCAAGATGGCACACCATGGCCAGGAAACAATACCGAGACC	1806
Qy	1551	ACCCGCGATGATCAGGTGTTCTTGGGCGCAGCGCGGCCACGACACCGAGGGCAACG	1610
Db	1807	ATCCTGGAATGATTACAGGTTTCTCTTGTGTCAAGTGTGGCTTGATATCGAGGCAATG	1866
Qy	1611	AGCTGCCCGCTCGTGTACGTTCTCCGTGAGAGCGCCCGGATTTCCAGCACCAAGA	1670
Db	1867	AGCTACCCCGTTTGGTCTATGTTTCTCGTGAAGAGCGCTCTGGATTTCCAGCATCAAGA	1926
Qy	1671	AGCCCGCGCCATGAACGCTCTGATTCGCGGTCTCGCGGTCTGACCAACCGCGCATTTCA	1730
Db	1927	AAGCTGGTGCATGAATGCTCTTGTGTCTCAGCTGTCTTACCAATGCAACAATACA	1986
Qy	1731	TGCTCAACTTGGACTGTGATCACTACATCAACACAGCAGGCGCATCCGGGAGCCATGT	1790
Db	1987	TGTTGAATCTTGTATGTGATCACTACATTAACAACAGTAAGGCTCTCAGGGAAGCTATGT	2046
Qy	1791	GCTTCTCATGAGACCTCAGGTCCGCGGAGGTCTGCTACGTTCAAGTTCAGTTCGCGAGAGGT	1850
Db	2047	GCTTCTTATGGAACCTTAACTAGGAAGGAGTGTCTGCTACGTCCAGTTTCCCCAGAGAT	2106
Qy	1851	TCGACGGCATGCAAGTGCACGACCGATACGCTAAACAGAAACACCGTCTTCTTCGACATCA	1910
Db	2107	TCGATGCAATTTGACAGGAATGATCGATATGCAACAGGAAACACCGTGTCTTCGATATTA	2166
Qy	1911	ACATGAAGGGCTCGAGCGGCATCCAAGCCCGGTGTAGCTCGGACAGGGTGGTGTTC	1970
Db	2167	ACTTGAGAGGTCTTTGATGGCATCCAAGAACAGTTTATGTGGAACCTGGCTGTGTTC	2226
Qy	1971	GGCGCCAGGGCTCTACGGCTTACAAACCTCTCCAAAGGACCCCAAGA-----	2015
Db	2227	ACCGAACAGCTCTATATGTTTATGAGCCCCCAATTAAAGCAGAAAGGGTGTCTTGT	2286
Qy	2016	-----GGCCCCAAGATGTGACCTGCGACTGTGCTGCCCTGTGCTCGGCGCAAGA	2063
Db	2287	CATCACTATGTGGCGGTAGGAAGAGCAAGCAAAATCAAAGAGGGCTCGCAACAAGA	2346
Qy	2064	AGCGGAACACGCCAAGGAC-----GGGCTGCCGGAGG	2096
Db	2347	AGTCGCAAGAGCATGTGGACAGTTCTGTGCCAGTATTCAACCTTGAAGATATAGAGGAGG	2406
Qy	2097	GCACCGCTGATATCGGAGTAGATAGCGACAGGAGATGCTCATGTCCACATGAACCTTCG	2156
Db	2407	GAGTTGAAGCGCTGGATTTGACGACGAGAAATCACTTCTATGTCTCAATAGCGCTGG	2466
Qy	2157	AGAAAGCGTTGCGGACGTCCGCGCGTTCGTCAAGTGCAGCTGATGAGAGGAGCGCGG	2216
Db	2467	AGAAAGATTGGCCAGTCCGCGAGCTTGTGCTCCACTCTCTGATGGAGTATGGTGTG	2526
Qy	2217	TCCCTCTCTGTCAGCGCCCGCGCTCTCAAGGAGGCGCATCCATGCTCATAGCTGCG	2276
Db	2527	TTCTCTAGTCCGCAACTCCGAGTCTCTTCTGAAGAAGTATCCATGTTTAAAGCTGTG	2586
Qy	2277	GCTACGAGGACAAAGACCGACTGGGGGCTGGAGCTGGGGTGGATCTACGGGTGATCAAGG	2336
Db	2587	GCTATGAGGACAAAGCTGGAATGGGGAATGAGATCGGGTGGATCTACGGTTCTGTGACAG	2646
Qy	2337	AGGACATCTGACGGGGTTCAAGATGCACTGCGCGGGTGGCGCTCCGTGTACTGCAATGC	2396
Db	2647	AAGACATTTCTACCGGATTCAGATGCAACGCGGAGGCTGGCGGTGATCTACTGCAATGC	2706
Qy	2397	CGAAGCGGCGCGCTTCAAGGGGTGCGGCGCGATCAATCTATCGGACGCTCTCAACGAGG	2456
Db	2707	CCAAAGCCGAGCTTTCAAGGGGTCTGCCCCCATCAATCTTTTCGGAACGCTGTGAACGAGG	2766
Qy	2457	TGCTCCGCTGGCGCTCGGGTCTCGAGATCTTCTTTCAGCGGCGCACAGCCCTGCTGT	2516
Db	2767	TGCTCCGCTGGGCTCTTGGGTCTCGTGGAGATCTCTTTCAGCGCGGACCTGCCCCCTGTGT	2826
Qy	2517	ACGCTTCAAGAAACCGCAACCTCAAGTGGCTGGAGCGCTTCGCCTACATCAACACCA	2576
Db	2827	ACGGCTAC--GGAGGGCGGCTCAAGTTCTCTGGAGAGATTCCGCTATCAACACCA	2883
Qy	2577	TCTACCCCTTCACTCTCCCTGCTCGCTACTGCACTGCACTCCCGCTGCTGCTGCTCC	2636
Db	2884	TCTACCCGCTCAGCTCCATCCCGCTTCTCATCTACTGATCTCTGCGGCACTGTCTGCTG	2943
Qy	2637	TCACCGCAAGTTCATCATGCGCTGATAGACGTTTCGCGAGGCTCTTCTTCTCATGCGCC	2696
Db	2944	TCACCGAAAGTTTCAATTCACAGAGATCAGCAACTTCGCCAGCATCTGGTTCATCTCC	3003

QY 1131 TCGTCAGCAGCGTGGACCCGCTCAAGGAGCGCCGCTGGTGACCGCCCAACCCGCTCT 1190
DB 1387 TCGTCAGTACAGTCGACCCCAATGAAGGAGCCCTCTTGTCTACCTGCCAATACCGTGTAT 1446
QY 1191 CCATCCTCGCGTAGACTACCCCGGTGGACAAGGTCTCTGCTTACGCTCCGACGACGGCG 1250
DB 1447 CCAITCTTGTCTGTGATTAACCTGTGGATAGGTCTCTTGTCTATCTGATCTGATGGAG 1506
QY 1251 CGTCGATCGTACGCTTCAGTCTGTGTGGAGACGGCCGAGTTCGCGCGCAAGTGGGTGC 1310
DB 1507 CTGCGATGCTGACATTTGATGCACTAGCTGAGACTTCAGAGTTTCTAGAAAAATGGGTAC 1566
QY 1311 CCTTCTGCAAGAAGTTTCGGCATCGAGCCCGCGCCCGGAGTTCTACTTCTCGTCTCAGG 1370
DB 1567 CATTTGTTAAGAAGTACCAACATTTGAACCTTAGAGCTCTCTGAATGGTACTTCTCCAGAAAA 1626
QY 1371 TCGACTACTCTCAAGGACAAGGTGACGCCACCTTCGTGCGAGAGCGCGCCGCTCAAGA 1430
DB 1627 TTGATTACTTGAAGGACAAGGTGACCCCTCATTTGTTAAGACCGCCGGGCCATGAAGA 1686
QY 1431 GAGAGTATGAGAGTTCAAGTTCGGATCAACGCGCTGGTGGCCAAAGCCCATGAAGTGC 1490
DB 1687 GAGAATATGAAGAAATCAAGATTAGGGTAAATGGCTTGTTCCTAAGGCACAGAAAGTTC 1746
QY 1491 CGGACAGGGGTGGATCATGAAGACGGCACGCCGTGGCCCGGGNAACAACCCCGGACC 1550
DB 1747 CTGAGGAAGGATGGATCATGAAGATGGCACACCATGGCCAGGAAACAATACGAGGACC 1806
QY 1551 ACCCGGATGATCATCAGGTGTTCTGGGCCACAGCGCGGCCACGACACGAGGSCAACG 1610
DB 1807 ATCTGGAAATGATTCAGGTTTCTTGTGTCAAGTGTGGCTTGTATCTGAGGGCAATG 1866
QY 1611 AGTGTCCCGCTCGTGTACGCTTCCGTGAGAAAGCGCCCGGATTCAGCACCAACAAGA 1670
DB 1867 AGCTACCCCGTTGGTCTATGTTCTCTGTAAGAGCGCTCTGGATTTCAGCATCACAGA 1926
QY 1671 AGGCGGCGCATGACGCTCTGATTCGGTCTCGCGCTGCTGACCAACGCGGCATTCA 1730
DB 1927 AAGCTGGTGCATGAATGCTCTTGTGTGTCTCAGCTGTGCTTACCAATGGACAATACA 1986
QY 1731 TGCTCAACTTGGACTGTGATCACTACATCAACAACAGCAAGCCCATCGGAGGCGCATGT 1790
DB 1987 TGTGAACTTGTATTGTGATCACTACATTAACACAGTAAAGCTCTAGGGAAAGCTATGT 2046
QY 1791 GCTTCTCATGAGCCCTCAGTTCGGCCGGAAGGTCTGCTAGTTCAAGTTCCTCCGAGAGGT 1850
DB 2047 GCTTCTTATGGAACCTTAACCTAGGAAGAGTGTCTGCTACGTCCTCAGTTTCCCGAGAGAT 2106
QY 1851 TCGACGGCATCGAGTGCACGACCGATAGCTTAACAGAAACACCGTCTTCTTCGACATCA 1910
DB 2107 TCGATGGCATTTGACAGGAATGATCGATATGCCAACAGGAACCCGTTGTTTTTCGATATTA 2166
QY 1911 ACATGAAGGGCTGGACGGCATCAAGGCCGCTGTAGCTCGGACAGGGTGGCTGTTC 1970
DB 2167 ACTTGAGAGGTCTTGATGGCATCCAGAACCAAGTTTATGTGCGAACTGGCTGTGTTTCA 2226
QY 1971 GGCGCCAGGCGCTCTACGGCTACAACTCCCAAGGGAACCAAGA----- 2015
DB 2227 ACCGAACAGCTCTATATGTTATGAGCCCCCAATTAAGCAGAAGAGGTTGTTCTTGT 2286
QY 2016 -----GGCCCAAGATGTGACTCGGACTGTGCGGCTGCTCGGCGCGCAAGA 2063
DB 2287 CATCACTATGTGGCGGTAGGAAGAAGGCAAGCAATCAAGAAAGGGCTCGGACAAAGA 2346
QY 2064 AGCGGAACACGCCAAGAC-----GGGCTGCGGAGG 2096
DB 2347 AGTCGAGAACGATGTGACAGATTCTGTGCCAGTATTCACTTGAAGATATAGAGGAGG 2406
QY 2097 GCACCGCTGATATGGAGTAGATAGCGACAAGAGAGATGCTCATGTGCCACATGAACATTTCG 2156
DB 2407 GAGTTGAAGGCGCTGGATTTGACGACGAGAAATCACTTCTTATGTCTCAATATGAGCCTGG 2466
QY 2157 AGAAGCGGTTCCGGGAGTTCGCGGGGTTTCGTACGTCGACGCTGATGGAGGAAGCGGGCG 2216

DB 2467 AGAAGAGATTTGGCCAGTCCGACAGGTTTGTGTGCTTCCACTCTGATGAGTATGGTGGTG 2526
QY 2217 TCCCTCTCTTCGTGAGCCCCCGCGCTCTCAAGGAGGCCATCCATCTCATCTCACTCGG 2276
DB 2527 TTCTCAGTCCGCAACTCCGAGTCTCTTCTGAAAGAAGCTATCCATGTTATAGCTGTG 2586
QY 2277 GCTACGAGGACAAGACCGACCTGGGGGCTGGAGCTGGGGTGGATCTACGGGTGATCAAG 2336
DB 2587 GCTATGAGGACAAGACTGAATGGGGAATGAGATCGGGTGGATCTTACGGTTCTGTGACAG 2646
QY 2337 AGGACATCTCGACGGGGTTCAAGATGACATGCGCGGGGTGGGCTCCGCTGACTGCATGC 2396
DB 2647 AAGACATTTCTCACCGGATTTCAAGATGACGCGGAGGCTGGCGGTGATCTACTGCATGC 2706
QY 2397 CGAAGCGGGCGCGCTTCAAGGGGTTCGCGCGCGATCAATCTATCGGACCGTCTCAACACG 2456
DB 2707 CCAAGCGGCAAGCTTTCAAGGGGTCTGCCCCCATCACTTTTCGACCGCTCTGAACACG 2766
QY 2457 TGCTCCGGTGGCGCTGGGGTCCGTGAGATCTTCTTCAGCGGGCAAGCCCTCTGTGT 2516
DB 2767 TGCTCCGGTGGGCTCTTGGGTCCGTGGAGATCTCTTTCAGCGGCACTGSCCCTGTGT 2826
QY 2517 ACGGCTACAGAAACGCGCAACCTCAAGTGGCTGGAGCGCTTCGCTACATCAACACCA 2576
DB 2827 ACGGCTAC--GGAGGGGCGGCTCAAGTTCTTGGAGAGATTCGCGTACATCAACACCA 2883
QY 2577 TCTACCCCTTCACTCGCTCCGCTGCTGCGCTACTGTCACCCCTCCCGCGCTCTGCTCTCC 2636
DB 2884 TCTACCCGCTCACTGCTCCATCCGCTTCTCATCTACTGATCTGCGCGCATCTGTCTGC 2943
QY 2637 TCACCGGCAAGTTCAATCATGCTCCGTGATAGCAAGTTTCGCGAGCTCTTTCATTCGCCC 2696
DB 2944 TCACCGGAAAGTTCAATCATTCAGAGATCAGCAACTTCGCGAGCATCTGTGTTCTCTCC 3003
QY 2697 TCTTATGTTCCATCTTCCGACGGGCTCTGGAGATCGGTGAGGCGGGTGGGATCG 2756
DB 3004 TCTTATGTTCCATCTTCCGACGGGCTCTGGAGATCGGTGAGGCGGGTGGGATCG 3063
QY 2757 AGGAGTGTGGAGGAACGAGCAGTTCTGGGTTCATCGCGGGCTGTCCGCGCATCTCTTCG 2816
DB 3064 ACAGTGTGGAGGAACGAGCAGTTCTGGGTGATCGGGGCTCTCCGCGCATCTCTTCG 3123
QY 2817 CCGTGTGCGAGGCTCTGCTAAGTCTCTGCGCGGATCGACCAACTTCAAGTCACT 2876
DB 3124 CCGTGTTCAGGGCTCTGCTCAAGGTGCTGCGCGGATCGACCAACTTCAAGTCACT 3183
QY 2877 CCAAGGCAACGCGGACGAGGACGAGTTTCGCGGAGCTTACGCTTCAAGTGGACCA 2936
DB 3184 CCAAGGCTCG--GACGAGGACGGGCTTTCGCGGAGCTGTACATGTTCAAGTGGACCA 3240
QY 2937 CGCTCTCTATCCCGCCCAACGAGTCTCATTAACGTCATCGGGTCTGTCGCGGCA 2996
DB 3241 CGCTCTGATCCCGCCCAACGAGTCTCATTAACGTCATCAAGTGGTGGGCTGTCGCGGCA 3300
QY 2997 TCTCCGACGCCATCAACAAACGAGTACAGTCTTGGGGGCCCTCTTCCGCAAGCTCTTCT 3056
DB 3301 TCTCTACGCCATCAACAGCGGATACAGTCTGCGGGGCCGCTCTTCCGCAAGCTCTTCT 3360
QY 3057 TCGGCTTCTGGTTCATGTCACCTCTACCGGTCCTCAAGGGGCTCATTGGGGGCCGCA 3116
DB 3361 TCGGCTTCTGGTTCATGTCACCTGTCACCGTTCCTCAAGGGGCTCATTGGGGGCCGCA 3420
QY 3117 ACAGGACGCCCAACCGGTTGTTGTCATCTGGTCTCATTTCTGGGCTCTCATCTTCTCCCTGC 3176
DB 3421 ACCGCAACCGCAACCATGTCGTCTGGGGCATCTCTGTTGGGCTCATCTTCTCTCTTCG 3480
QY 3177 TCTGGGTGAGATCGACCCCTTTCATGTCGAGGACCAAGGGGCCGAGCTCAGGAGTGTG 3236
DB 3481 TGTGGGTTCGATCGACCCCTTTCACCCCGGCTCACTGGGCCGAGTACCCAGAGCTGTG 3540
QY 3237 GCATCAATTTGCT 3248
|||||


```
Db      3541 GCATCAACTGCT 3552

RESULT 13
US-10-209-059-9
; Sequence 9, Application US/10209059
; Publication No. US20030163838A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Haiyin
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 3773
; TYPE: DNA
; ORGANISM: Zea mays
US-10-209-059-9

Query Match      40.68; Score 1397.6; DB 16; Length 3773;
Best Local Similarity 67.2; Pred. No. 0;
Matches 2185; Conservative 0; Mismatches 924; Indels 143; Gaps 9;

QY      133 GAACGAGCTGCTGCTGATCCGGGCGCCACGAGGACCCCAAGCCGCTGGGGCGCTGAGCGG 192
Db      325 GAGCTCGCTTGCCATGAGGGCGACGCGGACGGGTGAGTTCGGGAGGCGCGGTGGCGG 384
QY      193 GCAGGTGTGCGAGATATGCGCGGACGAGGTTCGGGCTCACGGTGGACGGACCTCTTCGT 252
Db      385 ACAGGTGTGCGAGATCTGCGGCGACGCGGTGGGCGACCAACGCGGAGGGGACGCTTCGC 444
QY      253 CGCTCTCAACGAGTGGCGCTTCCCGGTGTCGGGCGCTGTACGAGTACGAGCGCGGGA 312
Db      445 CGCTCTGCGAGCTGTGCGGGTTTCGGGTGTGCGGCGCTGTACGAGTACGAGCGCAAGGA 504
QY      313 GGGCAGCGCAAACTGCCGCCAGTGCAGACGCGCTACAAGCGCTCAAGGGGAGCGCGAG 372
Db      505 CGGCACGCGCGGTGCCCGCCAGTGCAGACCAAGTACAGCGGCCCAAGGGGAGCGCGC 564
QY      373 GGTTCGCGGGACGATGACGAGGAGACATCGACGACCTCGAGACGAGTTCACATCGA 432
Db      565 GATCCGTGGGGAGGAGGAGACGACACTGATGCCGATAGCGACTTCAATTACCTTGCATC 624
QY      433 CGACGGAATCAGCAGAGGAGCTGGAGGCGCAATCGACGCAACAGCCAGATCA----- 485
Db      625 TGGCAATGAGGACGAGAGCAGAGATTGCGGACAGAAATCGCAGATCGCGCTGGCGCAAGCGT 684
QY      486 -----CCGAGGCGATCTGCACGCGAGGATGAGCTACGGGAGGGGCGCCG----- 530
Db      685 TGGGGGAGCGGGGATGTTGGTGCCTCCCAAGTATGACAGTGGCGAGATCGGGCTTACCAA 744
QY      531 --ACGACGGCGAGCGGCAACACACCCCGCAGATCCCGCCCATCATCACCGGCTCCCGCTC 588
Db      745 GTATGACAGTGGCGAGATTCTCGGGGATACATCCCATCATGCTCACTAACAGCCAGATCTC 804
QY      589 CGTCCCGGTGAGCGGTGAGTTTCC-GATTACCAACGGGTATGGCCACGGCGAGGTCTCGT 647
Db      805 AGGAGAAATCCCTGGTCTTCCCTGACCATCATATGATGTCCCACTCGGGAACATTTGG 864
QY      648 CTTCCCTGCAACAGCGCATCATCCGTACCTGTGTCTGAGCCAGGAGTGCACAAAGTGGG 707
Db      865 CAAGCGTGTCTCCATTTC-CTATGTGAACCAATTCCGCCAAATCCGTCAAGGGAGTCTCTG 923

108 ACAGAGAAAGAAAGTGAAGTGAAGAGGAGGATGAGACGACTGGAAGTCAAGGAGGCA 767
124 GTAGCAATTGGGAATGTTGCTCGAAAGAGAGGGTTCATGGCTGGAATAATGAAGCAGGACA 983
168 -----TCC 770
184 AGGGGACGATTCATATGACAAATGGCAAGCAATGCTCCCTCTGAGGGTTCGGGGTGTG 1043
171 TCGGCGCGCGCGCGCATCCGAAGACATGGAAGACCGCGCGTGGCACTGAACGACGAGGCGCA 830
1044 GTGATATTGATGATCAATCAATGATTAACAATGAAAGATGCTTATTTGAACGACGAACTC 1103
831 GGCAGCGCGTGTGCGAGGAAGGTGTCGATCGCGTGCAGAGAGGTGAACCCCTACCGGATGG 890
1104 GACAGCTCTCATCTAGGAAAGTTCACCTTCCTCCAGGATAAATCCATACAGAGTGG 1163
891 TGATCTGTGTGCTCTCGTGTGCTCGCTTCCTTCCTCCGGTACCGGTATCCTGCACCCCG 950
1164 TCATTGTGTGCGATTTGATTTGTTAAGCATCTTCTTGCACTACCGGTATCACAATCTCTG 1223
951 TCCCGACGCGCATTCGGGCTGTGGCTCTCTCCATCATCTGCGAGATCTGGTTCGCCATCT 1010
1224 TGGCAATGATATACCATTAATGCTTCTATCTGTTATATGAGATCTGGTTCCTCTT 1283
1011 CTTGATCTCTGACAGTTCCTCCCAAGTGTTCCTCCATGACCGCGAGACGTACTCGAC 1070
1284 CGTGGATATTGGATCAGTTCCTTAAGTGGTTTCCAATCAACCGGAGAGACTACTTGATA 1343
1071 GCCTCTCCCTCAGGTACGAGAGGAGGGGCGGTGCTGCTGCTGCGGCTGGAACCTGT 1130
1344 GGTGTCATTAAGGTATGACCGGGAAGGTGAGCATCTCAGTTGGCTGCTGTTGACATTT 1403
1131 TCGTGAACAGCGGTGAGACCGCTCAAGGAGCGCGCTGCTGAGCGGCAACACCGTGTCT 1190
1404 TCGTCACTAGTACGACGACCAATGAAGAGCGCTCTCTGTCATGCGCAATACCGTGTAT 1463
1191 CCATCTCGCGGTAGACTACCGGTGAGCAAGGTCTCTGCTACGTCTCGACGAGCGG 1350
1464 CCATCTCTGTGTGGATTACCTGTGGATAAGGTCTCTTGTATGTATCTGATGATGAG 1523
1251 CGTGTGATGTGAGTTCGAGTCTGCTCGGAGAGCGCGGAGTTCGCGCGCAAGTGGTGC 1310
1524 CTGCGATGCTGACATTTGATGACACTAGCTGAGACTTTCAGAGTTTGTAGAAAATGGGTAC 1583
1311 CCTTCTGCAAGAGTTTCGCGCATCGAGCGCGCGCGCGGAGTTCTACTTCTCGCTCAAGG 1370
1584 CATTTGTTAAGAGTACAACTTGAACCTAGAGTCTCTGAAATGCTTCTCCAGBAAA 1643
1371 TCGACTACCTCAAGGACAGGTGTCAGCCCACTTCGTGCGAGGAGCGCGCGCATGAAGA 1430
1644 TTGATTTACTTGAAGGACAAAGTGCACCTTTCATTTGTTAAGACCGCGCGGCGCATGAAGA 1703
1431 GAGAGTATGAGGAGTTCAAGGTCCGATCAACCGCTGTTGGCGCAAGGCGCATGAAGTGC 1490
1704 GAGAAATGAGAAATTTCAAGTTAGGGTAAATGGCTTGTGCTAAGGACACAGAAAGTTC 1763
1491 CGGACAGAGGGTGGATCATGAAGGACGCGCGCTGGCGCGGGAACAAACACCGCGGAC 1550
1764 CTGAGGAGGATGATCATGCAAGATGGCACACCATGGCCAGGAAACAATACAGGAGC 1823
1551 ACCCGCGCATGATTCAGGTGTTCTCGGGCCACAGCGCGGCGCACGACACGAGGGGCAAG 1610
1824 ATCCTGGAATGATTCAGGTGTTCTTGGTGCACAGTGTGGCTTGTACTGAGGGGCAATG 1883
1611 AGCTGCGCGCTCGGTGATGCTCCCGTGAAGGCGCGCGGATTCACAGCACACAGA 1670
1884 AGCTACCCGCTTGGTCTATGTTTCTCGTGAAGAGCGTCTCTGATTCAGCATCACAGA 1943
1671 AGCGCGCGCATGAAAGCTCTGATTCGGGTCTCCCGCGGTCTGCTGACCAACGCGCCATTC 1730
1944 AAGCTGTCATGATGATGCTCTTGTGCTGCTCAGCTGTGCTTACCAATGAGCAATACA 2003
1731 TGCTCAACTTGGACTGTGATCACTACATCAACAGAGCAAGGCGCATCCGGAGGCGCATGT 1790
```

Db 2004 TGTGAATCTTGAATGATCACTATCAATTAACAACAGTAAGGCTCTCAGGAAGCTATGT 2063
Qy GCTTCTCATGACCCCTCAGGTCCGCGGAAAGGTCTGCTAGCTTTCAGTTCCTCCGAGAGGT 1850
Db GCTTCTTATGAGCCCTAACTAGGAAGAGTGTCTGCTAGCTCCAGTTCCTCCAGAGAT 2123
Qy TCGAGCGCATCGAGTGCAGACGATAGCTTAACAGGAACAACGCTCTTCTTCGACATCA 1910
Db TCGATGGCAATTCACAGGAATGATCGATATGCCAACAGGAACAACGCTGTTTTTCGATATTA 2183
Qy ACATGAAGGGCTGAGCGGCATCAAGGCCGCTGCTAGCTCGGACAGGGTGCCTGTTCC 1970
Db ACTTGAGAGTCTTGATGGCAATCCAGGACCAAGTTTATGTCGAACTGGCTGTGTTTTCA 2243
Qy GGCGCCAGGCGCTCTACGGCTACAACTCCCAAGGACCCAAAGA----- 2015
Db ACCGAACAGCTCTATATGTTATGAGCCCCCAATTAAGCAGAGAAGAGGGTGTCTTTGT 2303
Qy -----GGCCCAAGATGTTGACTCGGACTGCTGCCGCTGCTTCGGCGCAGAGA 2063
Db CATCATATGTTGGCGTATGGAAGAGGCAAGCAATCAAGAGAGGGCTCGGACAGAAGA 2363
Qy AGCGGAACACGCCAAGGAC-----GGGCTGCCGAGG 2096
Db AGTCGAGACATGTTGACACAGTTCTGTGCCAGTATTCAACTTGAAGATATAGAGGAG 2423
Qy GCACCGCTGATATGGGAGTAGATAGCGACAAGAGGATGCTCATGTCCACATGAACCTTCG 2156
Db GAGTTGAAGCGCTGGATTGACGACGAGAAATCACTTCTTATGCTCAATAGACCTGG 2483
Qy AGAAGCGTTCGGGAGTCCGGCGGTTTCGTCAGTCGACGCTGATGAGGAAGAGCGCG 2216
Db AGAAGAGATTTGGCGAGTCCGCGAGTTTGTGTGCTCCACTCTGATGGAGTATGGTGTG 2543
Qy TCCCTCTCTGTCGAGCCCCCGCGCTCTCAAGGAGGCCATCATGTCATCAGCTGCG 2276
Db TTCCTCAGTCGCAACTCCGAGTCTCTTCTGAAGAAGCTATCATATGTTATAGCTGTG 2603
Qy GCTACGAGGACAAGACCTGAGCTGGGGCTGAGCTGGGTGGATCTACGGGTCTGATCAAG 2336
Db GCTATGAGGACAAGACTGAATGGGGAACCTGAGATCGGGTGGATCTACGGTCTGTGACAG 2663
Qy AGGACATCTGACGGGTTCAAGATGCACTCCGCGGGTGGCTCCGCTGATGCTGATGCTG 2396
Db AAGACATCTCACCGGATTCAGATGCAAGGCGGAGGCTGGCGGTCTGATCTACTGCTATG 2723
Qy CGAAGCGGCGGCTTCAAGGGTTCGGCGGATCAATCTATCGAAGCGCTCTCAACGAG 2456
Db CCAAGCGGCCAGCTTCAAGGGGTCTGCCCCCATCAATCTTTCGAGCGCTGTAACACCA 2783
Qy TGCTCCGGTGGCGCTGGGGTCCGTCGAGATCTTCTTCAGCGGCGCACAGCCCTGCTGT 2516
Db TGCTCCGGTGGCTCTTGGTCCGTCGAGATCTTCTTCAGCGGCACTGCCCTGCTGTGT 2843
Qy ACGGCTACAGAAACGCAACCTCAAGTGGCTGAGCGCTTCGCTTACATCAACACCA 2576
Db ACGGCTAC--GAGAGGGCGGCTCAAGTTCCTGGAGAGATTCGCGTATCAACACCA 2900
Qy TCTACCCCTTCACTCGCTCCGCTGCTCGCTACTGACCTCCCGCGCTCTGCTCC 2636
Db TCTACCCGCTCAGCTTCATCCCGCTTCTCATCTACTGATCTTCCCGGCACTGTCTGC 2960
Qy TCACCGGCAAGTTTCAATCATGCTGCTGATTCAGCGCTTTCGCGAGCTCTTCTTCACTCGCC 2696
Db TCACCGGAAGTTTCAATCTCCAGATACAGCACTTCGCGAGATCTGTTTCACTCTCC 3020
Qy TCTTCACTCATCTTCGCGACGGGCACTCTTGGAGATGCGGTGAGCGGGTGAAGCTG 2756
Db TCTTCACTCATCTTCCGCAAGGCACTCTTGGAGATGCGGTGAGCGGGTGAAGCTG 3080
Qy AGGAGTGGTAGGAACAGGACAGTCTTGGGTTCATCGCGGCGGTCTCCGCGCATCTCTTCG 2816

Db 3081 ACGAGTGGTAGGAACAGCAGGATTTCTGGGGTATCGGGGGGATCTTCGCGGCACTCTTCG 3140
Qy CCGTCTGTCAGGGGCTGCTCAAGGTCTTCGCGGGATCGACACCAACTTTCACCGTCACT 2876
Db CCGTGTTCAGGGGCTGCTCAAGGTCTTCGCGGGATCGACACCAACTTTCACCGTCACT 3200
Qy CCAAGGCCACCGGCGAGGACGACGAGTTTCGCGGAGCTTACATGTTCAAGTGGACGA 2936
Db CCAAGGGCTCG--GACGAGGACGCGGACTTCGCGGAGCTGTACATGTTCAAGTGGACGA 3257
Qy CGCTCTCATCCTCCGCCCAACAGCGTGTCTCATATTAAGCTCATCGGGTCTGCGCGGCA 2996
Db CGCTCTCATGATCCGCCCAACAGCACTGTATCATCAACTGCTGCGGCTGCTGCGCGGCA 3317
Qy TCTCCGACGCGCATCAACACGCGGTACAGTCTCTGGGGGCCCTTTCGCGCAAGCTCTTCT 3056
Db TCTCTACGCGCATCAACAGCGGATACAGTCTGCGGGGCCCTTTCGCGCAAGCTCTTCT 3377
Qy TCGCTTCTGGGTATGCTCCACCTTACCGGTTCTCAAGGGGCTCATGCGGGCGGCA 3116
Db TCGCTTCTGGGTATGCTCCACCTTACCGGTTCTCAAGGGGCTCATGCGGGCGGCA 3437
Qy ACAGACGCGCCACCGTGTGTTGTCATCTGCTGCTCATTTCTGCTGCGCTTCATCTTCTCCCTGC 3176
Db ACAGACGCGCGACCATGCTGCTGCTGCGGCACTCTGCTGCGGCTTCATCTTCTTCTTCTTCT 3497
Qy TCTGGGTGAGGATACGACCTTTCATGCTCAGGACCAAGGGGCCGAGCTCAGGAGTGTG 3236
Db TGTGGGTTCGATCGACCCCTTTCACCAACCGCGTCACTGGCGCGGATACCCAGAGCTGTG 3557
Qy GCATCAATTGCT 3248
Db GCATCAACTGCT 3569

RESULT 14
US-10-160-719-5
; Sequence 5, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 3773
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (338) ... (3568)
US-10-160-719-5

Query Match 40.6%; Score 1397.6; DB 16; Length 3773;
Best Local Similarity 67.2%; Pred. No. 0;
Matches 2185; Conservative 0; Mismatches 924; Indels 143; Gaps 9;
Qy 133 GAACGAGCTGCTGCTATCGGGGCCACGAGGACCCCAAGCGCTGCGGGCGCTGAGCGG 192
Db 325 GAGCTCGCTTGGCCATGGAGGCGGACGCGGTGAAGTTCGGGAGGCGCGGTGGCGG 384
Qy 193 GCAGGTGTGAGATATGCGGCGGAGGTCGGGCTCACGGTGGACCGGACCTCTTCGT 252

Db 385 ||||| ACAGGTGTCAGATCTGCGGCGACGCGTGGGCACCAACGCGCGAGGGGACGCTTCGC 444
Qy 253 CGCCTGCAACGAGTGCAGCTTCCCGGTGTCGCGCCCTGCTACGAGTACGAGCGCGCGGA 312
Db 445 CGCCTGCGAGCTGTCGCGGTTTCGGTGTGCGCGCCCTGCTACGAGTACGAGCGCAAGGA 504
Qy 313 GGGCACGACAGACTGCCCCAGTGCACAGCGCGCTACAAGCGCTCAAGGGGAGCCGAG 372
Db 505 CGGCACGCGAGCGTGCCTCCAGTGCACAGCAACAGTACAAGCGCCACAAAGGGGAGCCGCG 564
Qy 373 GGTTCGCGGGACGATACGAGGAGGACATGACGACCTGGAGCAAGGTTCAACATCGA 432
Db 565 GATCCGTGGGGAGGAAGGACGACACTGATGCCGATAGCGACTTCAATTACCTTGCATC 624
Qy 433 CGACGAGAACTACGACAGAGGAGCTGGAGGGCAACATGCAGAACAGCCAGACATCA----- 485
Db 625 TGGCAATGAGGACCAAGAACGAAAGATTGGCGACAGAAATGCGCAGCTGGCGGCATGAACGT 684
Qy 486 -----CCGAGGCGATCTGCACGCGCAGGATGAGCTACGGGAGGGGCCCG----- 530
Db 685 TGGGGGCGCGGGGATGTTGTCGCCCAAGTATGACAGTGGCGGAGATCGGGCTTACCAA 744
Qy 531 --ACGACGGCGACGGCAACACCCCGGAGATCCCGCCCATCATACCGGCTCCCGCTC 588
Db 745 GTATGACAGTGGCGGAGATTCTCGGGGATACATCCCATCAGTCACTAACACGCGAGATCTC 804
Qy 589 CGTCCCGGTGAGCGGTGAGTTTC-GATTACCAACGGGTATGGCCACGGCGAGGTCTCGT 647
Db 805 AGGAAATCCCTGGTCTTCCCTGACCATCATATGATGCCCAACTCGGAACATTTG 864
Qy 648 CTTTCCCTGCACAAAGCGCATCCATCCGTACCTGTGCTGACCCAGGAGTGCCAAAGTGG 707
Db 865 CAAGCGTCTCCATTTC-CTATGTGAACCATTCGCCAATCCGTCAAGGGAGTTCTCTG 923
Qy 708 ACGAAGAAGAAGTACGTGGAGGAGAGGATGGACGACTGGAAGTCCAAAGGCGCA 767
Db 924 GTAGCATTTGGGAATGTTGGCTGGAAAGAGAGGGTTGATGGCTGGAAATGAAGCAGGACA 983
Qy 768 -----TCC 770
Db 984 AGGGACGATTCCCATGACGANTGGCAACGATTGCTCCCTCTGAGGGTTCGGGGTGTG 1043
Qy 771 TCGGCGCGCGCGATCCCGAAGACATGGACGCGCAGCTGGCACTGAACGACGAGCGGA 830
Db 1044 GTGATATTGATGCATCAACTGATTACAACATGGAAGATGCTTATTGAACGACGAAACTC 1103
Qy 831 GGCAGCGCTGTCGAGGAAGGTTCGATCGGTGAGCAAGGTGAACCCGTACCGATGG 890
Db 1104 GACAGCCTCTATCTAGGAAAGTTCCACTTCTCTCCAGGATAAATCCATACAGGATGG 1163
Qy 891 TGATCGGTGCGTCTCGTTGTGTCGCCCTTCTTCTCCGTACCGTATCCTGCACCCCG 950
Db 1164 TCATTGTGTCGATTAATGTTCTAAGCATCTTCTTGCACTACCGTATCAAAATCCTG 1223
Qy 951 TCCCGGACGCCATCGGGCTGTGGCTCGTCTCCATCATCTCGAGATCTGGTTCGCCATCT 1010
Db 1224 TGGCAATGCATACCCATTATGGCTTCTATCTGTTATATGTGAGATCTGGTTGCTCTTT 1283
Qy 1011 CCTGGATCTCGACCACTTCCCAAGTGGTTCCCATCGACCGGAGACGTACTCGACC 1070
Db 1284 CGTGGATATTGGATTCAGTTCCCTAAGTGGTTTCCAATCAACCGGGAGACGTACTTGGATA 1343
Qy 1071 GCCTCTCCCTCAGGTACGAGGGAAGGGAGCGCTCGCTGTGTCGGCGGTGACCTGT 1130
Db 1344 GGCTGGCATTAAGGTATGACCGGGAAGGTGAGCCATCTCAGTTGGCTGTGTTGACATTT 1403
Qy 1131 TCGTGACGCGTGGACCCGCTCAAGGAGCGCGCTGGTGACCGCAACACCGTGTCT 1190
Db 1404 TCGTCAGTACAGTCGACCCCAATGAAGGAGCTCTCTTGTCTCACTGCCCAATACCGTGTAT 1463
Qy 1191 CCATCCTCGCGTAGACTACCCGCTGACAAAGTCTCTGCTAGCTCTCCGACGACGGCG 1250

Db 1464 CCATTCTTGTGTGGATTACCTGTGGATAAAGGTCTCTTGTATGTATCTGTATGTATGAG 1523
Qy 1251 CGTCGATGTGACGTTTCGAGTTCGTCGAGACGCGCCGAGTTTCGGCGCGCAAGTGGGTGC 1310
Db 1524 CTCGGATGTGACATTTGATGCACTAGCTAGACTTCAGAGTTTGTAGAAATGGGTAC 1583
Qy 1311 CTTTCTGCAAGAAATTGGGATCGAGCCCGCGCCCGCGAGTTCTACTTCTCTCCTCAAGG 1370
Db 1584 CATTTGTTAAGAGTACAACATTTGAACCTAGAGCTCTCTGAATGGTACTTCTCTCCAGAAA 1643
Qy 1371 TCGACTACCTCAAGGACAAGGTGCGACCCACCTTCCTGTCAGGAGCGCGCGCCCATGAGA 1430
Db 1644 TTGATTACTTGAAGGACAAAGTGCACCTTCAATTTGTTAAAGACCCCGCGGCATGAAGA 1703
Qy 1431 GAGAGTATGAGGAGTTCAAGTCCCGATCAACCGCTGGTGGCAAGGCCCATGAAGTGC 1490
Db 1704 GAGAAATGAGAAATTCAAAGTTAGGTAAATGGCCTTGTCTAAGGCACAGAAAGTTC 1763
Qy 1491 CGCAGAGGGGTGGATCATGAAGGACGCAACGCGCTGGCCCGGGAAACAACCCCGCAAC 1550
Db 1764 CTGAGGAAGGATGGATCATGCAAGATGGCACACCATGGCCAGGAAACAATACCAGGACC 1823
Qy 1551 ACCCCGCGATGATCCAGGTGTTCTCGGGCACAAGCGCGCGCCACGACACGAGGGCAAG 1610
Db 1824 ATCCTGGAATGATTCAAGTGTTCCTTGGTTCACAGTGTGGCCTTGATCTAGAGGGCAATG 1883
Qy 1611 AGCTGCCCGCTCGGTGTACGTCTCCCGTGAGAAGCGCCGCGATTCACAGCACCAAGA 1670
Db 1884 AGCTACCCGTTTGGTCTATGTTTCTCGTGAAGAGGTCTTGGATTCAGCATCACAAGA 1943
Qy 1671 AGSCCGCGCCATGAACGCTCTGATTCGCTTCCCGCGTGTGTCGACCAACCGCCCATTC 1730
Db 1944 AAGCTGTGCCATGAATGCTCTTGTTCGTGCTCTCAGCTGTGCTTACCAATGGACAATA 2003
Qy 1731 TGCTCAACTTGGACTGTGATCACTACATCAACACAGCAAGGCATCCGGAGGCCATGT 1790
Db 2004 TGTGGAATCTTGATTTGTGATCACTACATTAACAACAGTAAGGCTCTCAGGGAAGCTATGT 2063
Qy 1791 GCTTCTCATGGACCCCTCAGGTCCGCGGAAGGTCTGTACGTTCAGTTCCCGCAGAGGT 1850
Db 2064 GCTTCTTATGACCCCTAACCTAGGAAGGATGTCTGTACGTCCAGTTTCCCGCAGAT 2123
Qy 1851 TCGACGGCATCGACGTGCGACGCGATACGTTAAACAGGAACACCGTCTTCTTCGACATCA 1910
Db 2124 TCGATGGCATTCAGAGGAATGATCGATATGCCAACAGGAACACCGTGTCTTCGATATTA 2183
Qy 1911 ACATGAAGGGCTGGAGCGCATCCAGGCCCGGTGTACGTTCGGGACAGGGTGTGTTCC 1970
Db 2184 ACTTGAGAGTCTTGATGGCATCCAGGACCAAGTTTATGTTCGGAACCTGGCTGTGTTTCA 2243
Qy 1971 GCGCGCACGGCGCTCTACGGCTACAACCCCTCCCAAGGGAACCAAGA----- 2015
Db 2244 ACCGAACAGCTCTATATGTTTATGAGCCGCCCAATTAAAGCAGAAGAGGGTGTCTTGT 2303
Qy 2016 -----GGCCCAAGATGGTACCTGCGACTGTGTCGCCGTCTTCGGCGCGCAAGA 2063
Db 2304 CATCACTATGTGGCGGTAGGAAGCAAGCAAAATCAAAAGAGGGCTCGACAAGAAGA 2363
Qy 2064 AGCGGAAAACGCCAAGGAC-----GGCTCCCGGAGG 2096
Db 2364 AGTCGAGAGCATGTGGACAGTTCTGTGCCAGTATTCAACTTGAAGATATAGAGAGG 2423
Qy 2097 GCACCGCTCATATGGGAGTAGATAGCGACAAGGAGATGCTCATGTCCCATGTAACCTTCG 2156
Db 2424 GAGTTGAAGGCGCTGATTTGACGACGAGAAATCACTTCTATGTCTCAATGAGCCTGG 2483
Qy 2157 AGAAGCGGTTCGGGCAAGTCCGCGCGTTCGTCACTGCGAGCTGTGAGGAGGAAGCGGCG 2216
Db 2484 AGAAGAGATTGGCCAGTCCGACAGCGTTTGTTCCTTCCCTCACTCTGATGGAGTATGGTGG 2543
Qy 2217 TCCCTCTTCTGTCGAGCGCCCGCGCTCTCAAGGAGGCCATCCATGTTCATCAGCTGCG 2276
Db 2544 TTCTCAGTCCGCAACTCCGAGTCTCTCTGAAAGAGCTATCCTATGTTTAAAGCTGTG 2603

Qy	2277	GCTACGAGGACAGACCGACTGGGGCTGGAGCTGGGGTGGATCTACGGGTGCATCACGG	2336
Db	2604	GCTATGAGGACAGACTGAATGGGAACTCAGATCGGGTGGATCTACGGTTCCTGTGACAG	2663
Qy	2337	AGGACATCTCTGA CGGGGTTCAAGATGCACTGCGCGGGTGGCGTCCGTGTACTGCATGC	2396
Db	2664	AAGACATTTCTCACCGGATTCAGATGCACGCGGAGGCTGCGGTGCATCTACTGTCATGC	2723
Qy	2397	CGAAGCGGGCGGTTCAAGGGGTGCGGCGCGATCAATCTATCGGACCGTCTCAACGAG	2456
Db	2724	CCAAGCGCCAGCTTTCAAGGGGTCTGCCCCCATCAATCTTTGCGACCGTCTGAACGAG	2783
Qy	2457	TGCTCCGGTGGGCGCTGGGCTCGGTGAGATCTTCTACGCGGCAAGGCCCTCTGCTGT	2516
Db	2784	TGCTCCGGTGGGCTCTTGGGTCTGTGGAGATCTCTTCAGCGGCACTGCCCTCTGTGT	2843
Qy	2517	ACGGCTACAAGAACGGCAACTCAAGTGGCTGGAGCGCTTCGGCTACATCAACACCA	2576
Db	2844	ACGGCTAC--GGAGGGCGGCTCAAGTTCCTGGAGAGATTCGGGTACATCAACACCA	2900
Qy	2577	TCTACCCCTTCACTCGCTCCGCTGCTCGCTACTGCAACCTCCCGCGCTGCGCTCC	2636
Db	2901	TCTACCCGCTCACTGTCATCCGCTTCTCATCTACTGCACTCTGCCCGCATCTGTCTG	2960
Qy	2637	TCACCGCAAGTTTCATCATCGCTCGATTAGCACGTTTCGCAGCGCTCTTCTTCATCGCC	2696
Db	2961	TCACCGCAAGTTTCATCATTCAGAGATCAGCAACTTCGCCAGCATCTGTTTCATCTCC	3020
Qy	2697	TCTTCATGTCATCTTCGCAACGGCATCTGAGATCGGGTGGAGGGGGTGGAGATCG	2756
Db	3021	TCTTTCATCTCGATCTTCGCAACGGCATCTGAGATGAGGTGGAGGGGGTGGGATCG	3080
Qy	2757	AGGAGTGGTGGAGGAACGACAGTTCTGGGTCACTCGCGCGGTGTCCGGGCATCTCTCG	2816
Db	3081	ACGAGTGGTGGAGAACGACAGTTCTGGGTGATCGGGGCACTTCGCGCACTCTTCG	3140
Qy	2817	CGTCTGTGAGGGCTGCTCAAGGTCTCGCGGGATCGACACAACTTCACCGTCACT	2876
Db	3141	CGGTGTTCAGGGCTGCTCAAGGTGTGCGCGGCATCGACACAACTTCACCGTCACT	3200
Qy	2877	CCAAGGCCACGGCGACGAGAGACGAGTTCCCGAGCTCTACGCTTCAAGTGACCA	2936
Db	3201	CGAAGGCTCG--GACGAGAACGGGCACTTCGCGAGCTGTACATGTTCAAGTGACCA	3257
Qy	2937	CGTCTCTATCCCGCCCAACGCTGCTCATCATTAACGTCATCGCGCTCGTGGCGGCA	2996
Db	3258	CGCTCTCTGATCCCGCCCAACCACTCTGATCATCAACTGGTGGCGCTCGTCCCGGCA	3317
Qy	2997	TCTCCGAGCGCATCAACAGGGTACAGTTCCTGGGGGCCCTTCCTGGCAAGCTCTCT	3056
Db	3318	TCTCTACGCCATCAACAGCGGATACAGTGTGGGGCCGCTCTTCGGCAAGCTCTCT	3377
Qy	3057	TCGCTCTTGGGTCACTCGTCCACCTCTACCCGTTCTCAAGGGGCTCATGGGCGGCA	3116
Db	3378	TCGCTTCTGGGTCACTGTCACCTGTACCCGTTCTCAAGGGCTCATGGGCGGCA	3437
Qy	3117	ACAGGACGCCACGGTTGTTCATCTGGTCCATCTGTGGGCTCCATCTTCCTCGC	3176
Db	3438	ACCGCAACCCGACCATGCTGCTGTGGGCCCATCTGTGGGTGTCATCTTCCTTCG	3497
Qy	3177	TCTGGGTGAGGATCGACCCCTTTCATCGTCAGGACCAAGGGCCCGGACGTCAGGCA	3236
Db	3498	TGTGGGTTCGATCGACCCCTTTCACCAACCGGCTCACTGGCCCGGATACCGACG	3557
Qy	3237	GCATCAATTGCT	3248
Db	3558	GCATCAACTGCT	3569

RESULT 15

RESOL 13
US-10-627-132-9

US-10-627-132-3
; Sequence 9, Application US/10627132

```

; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 3773
; TYPE: DNA
; ORGANISM: Zea mays
US-10-627-132-9.

```

Query Match 40.6%; Score 1397.6; DB 17; Length 3773;
Best Local Similarity 67.2%; Pred. No. 0;
Matches 2185; Conservative 0; Mismatches 924; Indels 143; Gaps 9;

133	Qy	GAACGAGCTGGTGTGTGATTCGGGGCCACAGAGGACCCCAAGCCGCTGCGGGCGCCTGAGCGG	192
325	Db	GAGCTCGCTTTGCCATGTGAGGGCGCAGCGGGACCGGGTGTGAAGTTCGGGGAGGCGCGTGGCGG	384
193	Qy	GCAGGTGTGGAGAGATATCGGCGCAGCAGGCTCGGGCTCACGCTGGACGCGACCTCTTCGT	252
385	Db	ACAGTGTGCCAGATCTGCGCGACCGGCTGGGCACACGCGGAGGGGAGCGTCTTCGC	444
253	Qy	CGCTGTCAACGAGTGTGGCTTCCCGTGTGCCGGCCCTGTCTACGAGTAGTCGAGCGCGGGA	312
445	Db	CGCTCGCAGCGTCTGCGGGTTTCGGGTGTGCCCGCCCTGCTACGAGTAGTCAGCGCGCAAGGA	504
313	Qy	GGCAGCGCAGNATGCCCCCAGTGCAGACGCGGTACAAAGCGCTCAAGGGGAGCCCGAG	372
505	Db	CGGCACGCAAGCGCTGCCCGCAGTGCAGACCAAGTACAAGCGCCACAAAGGGGAGCCCGGC	564
373	Qy	GGTTGCCGGGGACGATCAGCAGGAGGAGCATCGACGACCTGGAGCACAGGTTCAACATCGA	432
565	Db	GATCCGTGGGGAGGAGGACGACCTGATGCCGATAGCGACTTCAATTACCTTGCATC	624
433	Qy	CGACGAGAAATCAGCAGAGGCGAGCTGGAGGGCAACATGCAGAAACAGCCAGATCA-----	485
625	Db	TGGCAATGAGGACCCAGAAGCAGAGAATTTGCCGACAGAAATGCGCAGCTGCGCGCATGAACGT	684
486	Qy	-----CCGAGGGCGATGCTCACGGCAGGATGAGTACGGGAGGGGCCCG-----	530
685	Db	TGGGGGCGCGGGGATGTGTGTGCCCCCAAGTATGACAGTGGCGAGATCGGGCTTACCA	744
531	Qy	--ACGACGGCGACGGCAACAACACCCCGCAGATCCCGCCCATCATCACCGGCTCCCGCTC	588
745	Db	GTATGACAGTGGCGAGATTCCTCGGGGATACATCCCATCAGTCACTAACAGCCAGATCTC	804
589	Qy	CGTGCCGGTGAGCGGTGAGTTTCC-GAATTACCAAGGGTATGGCCACGGCGAGGTCTCGT	647
805	Db	AGGAGAAATCCCTGGTCTTCCCTGACCATCATATGATGTCCCAACTGGGAAACATTGG	864
648	Qy	CTTCCCTGCAACAGCGCATCATCCGTACCTGTGTCTGAGCCAGGGAGTGCCAAGTGGG	707
865	Db	CAAGCGTCTCCATTTC-CTATGTGAACCAATTCGCCAAATCCGTCAAGGGAGTTCTCTG	923
708	Qy	ACGAGAGAAGAAGTGAAGTGGAGGAGGATGGACGACTGGAAGTCCCAAGCAGGCA	767
924	Db	GTAGCATTTGGGAATGTTTCTCGAAGAGAGGGTTGATGGCTGGAATAATGAACAGGACA	983

QY 768 -----TCC 770
Db 984 AGGGACGATTTCCCATGACCAATGGCAAGCAATTTGCTCTCTGAGGGTCGGGGTTTG 1043
QY 771 TCGSCGGCGCGCGATCCGGAACACATGGACCGCGAGCTGGCACTGAACGACGAGCGGA 830
Db 1044 GTGATATTGATGCATCACTGATTTACACATGGAGATGCTTATTGAACGACGAACACTC 1103
QY 831 GGCAGCCGCTGTTCGAGGAAGGTTCGATCGCGTCGAGCAAGGTGAACCCCGTACCAGATGG 890
Db 1104 GACAGCCTCTATCTPAGGAAGTTCCACTTCTCTCCAGGATAAATCCATACAGATGG 1163
QY 891 TGATCGTGGTGGCTCTGTTGCTCGCTCGCTTCTTCCTCCGGTACCGTATCTGTGACCCCG 950
Db 1164 TCATTGCTGCGATTTGATTTTAAGCATCTTTCTTGCACTACCGTATCACAAATCCTG 1223
QY 951 TCCCGACGCCATCGGCTGGCTCGTCTCCATCATCTCGAGATCTGGTTCGCCATCT 1010
Db 1224 TGGCAATGCATACCCATTATGGCTTCTATCTGTATTATGTGAGATCTGGTTTGTCTTT 1283
QY 1011 CCTGATCTCGACCACTTCCCAGTGGTTCCTCATCGACCGCGAGACGTACCTCGACC 1070
Db 1284 CGTGGATATTGGATCAGTCCCTAAGTGGTTTCCAAATCAACCGGGAGACGTACCTTGATA 1343
QY 1071 GCCTCTCCCTCAGGTACGAGGGAAGGAGGAGCGCTGCTGCTGTGCGCGGTGGAACCTGT 1130
Db 1344 GGCTGGCATTTAAGGTATGACCGGAAGGTGAGCCATCTCAGTTGGCTGCTGTGACATTT 1403
QY 1131 TCGTGACAGGTGGACCGCTCAAGAGCGCGCTGTGACCGCCAAACACCGTGTCT 1190
Db 1404 TCGTCAGTACGTGACCCCAATGAGAGGCTCTCTTTGTCATCTGCCAATACCGTGTAT 1463
QY 1191 CCATCTCGCGGTAGACTACCCGTGGAACAAGTCTCTGCTAGCTCTCGAGCAAGCGCG 1250
Db 1464 CCATTCTGTGTGGATTACCTCTGGATAAGGTCTCTTGTATGATCTGATGATGAG 1523
QY 1251 CGTCGATGCTGAGTTCAGTTCGCTGCGAGAGCGCGAGTTCGCGCGCAAGTGGGTGC 1310
Db 1524 CTGCGATGCTGACATTTGATGCACTAGCTGAGACTTCAGAGTTTGTGAGAAAATGGGTAC 1583
QY 1311 CTTTCTCAAGAAGTTCGGCATCAGCCCGCGCGCGGAGTTCTACTCTCGCTCAAGG 1370
Db 1584 CATTTGTTAAGAGTACAACATTGAACTTAGAGTCTCTGATGGTACTTCTCCAGAAA 1643
QY 1371 TCGACTACTCAAGGAAGGTGAGCCCACTTCTGTCAGGAGCGCGCGCCATGAAGA 1430
Db 1644 TTGATTACTTGAAGGACAAAGTGCACTTCAITTTGTTAAAGACCGCGCGGCATGAAGA 1703
QY 1431 GAGATATGAGGATTCAGGTCCGATCAACGGCTGGTGGCCAGGCGCATGAGGTGC 1490
Db 1704 GAGAATATGAAGATTCAGGTAAATGGGCTTGTGCTAAGGCAACAGAAAGTTC 1763
QY 1491 CGGACAGGGGTGGATCATGAAGGACGGCACCGCTGCGCGGGAACAAACACCGCGACC 1550
Db 1764 CTGAGGAAGATGATCATGCAAGATGGCACACCATGGCCAGGAAACATACAGGGACC 1823
QY 1551 ACCCGGCATGATCAGGTGTTCTTGGGCCACAGCGCGCGCCAGCACACCGAGGGCAAG 1610
Db 1824 ATCTCGAATGATTCAGGTTTCTTGTGTCACAGTGGTGGCTTGTACTAGGGCAATG 1883
QY 1611 AGTGGCCCGCTCGTGTAGTCTCCCGTGAGAAAGCGCCGGATTCAGGACCAAGA 1670
Db 1884 AGCTACCCCGTTGGTCTATGTTTCTGTTGAAAGCGTCTCGGATTCAGCATCACAGA 1943
QY 1671 AGGCGGCGCATGAACGCTCTGATTCGCTCTCCGCTGCTGACCAACCGCGCATTTCA 1730
Db 1944 AAGCTGTTGCCATGAATGCTTGTGTTCTGCTGCTGCTGCTTACCAATGGACAATACA 2003
QY 1731 TGCTCAACTTTGGATGATCACTACATCAACACAGCAAGGCGCATCCGGGAGGCCATGT 1790
Db 2004 TGTGTAATCTTGAATTGTGATCACTACATTAAACACAGTAAGGCTCTCAGGGAGCTATGT 2063
QY 1791 GCTTCTCTGAGGCCCTCAGGTGGCGGGAAGGTCTGCTAGTTCAAGTTCAGTTCGCCGAGAGGT 1850

Db 2064 GCTTCTTATGGACCCCTAACTAGGAAGGAGTGTCTGTACGTCCAGTTCCTCCAGAGAT 2123
QY 1851 TCGACCGCATTCGACGTGACGACCGATAGCTAAACAGGAACACCGTCTTCTTCGACATCA 1910
Db 2124 TCGATGGCATGACAGAAATGATCGATATGCCAACAGGAACACCGTGTCTTTCGATATTA 2183
QY 1911 ACATGAAGGGCTTGGACGGCATCAAAGGGCCGCTGTACGTCCGGACAGGGTGGTGTTC 1970
Db 2184 ACTTGAGAGTCTTGATGGCATCCAAGGACCACTTATGTTCGAACTGGCTGTGTTTTCA 2243
QY 1971 GGGCCAGGGCTCTAGGGCTACAACCTCTCCGAAGGACCCAAAGA ----- 2015
Db 2244 ACCGAAACAGCTCTATATGGTTATGAGCCCCCAATTAAGCAGAAAGAGGGTGGTTCTTGT 2303
QY 2016 -----GGCCCAAGATGCTGACCTGCGACCTGCTGCGCGTCTTCGGCGGCAAGA 2063
Db 2304 CATCACTATGTGGCGGTAGGAAGAGCAAGCAAAATCAAAAGAGGGCTCGGACAGAAAGA 2363
QY 2064 AGCGAAACACGCCAAGGAC -----GGGCTGCCGAGG 2096
Db 2364 AGTCGCAAGACATGTGGACAGTTCTGTGCCAGTATTCAACCTTTGAAGATATAGAGGAGG 2423
QY 2097 GCACCGCTGATATGGAGTAGATAGCGACAGAGAGATGCTCATGTCCCATGAACTTCG 2156
Db 2424 GAGTTGAAGCGCTGGATTGTGACGACGAGAAATFCACTTCTTATGTCTCAAATGAGCCGTG 2483
QY 2157 AGAAGCGGTTTCGGGCGAGTCCGCGCGTTCGTACGCTCGACGCTGATGGAGAAAGCGGCG 2216
Db 2484 AGAAGAGATTTGGCCAGTCCGAGCGTTTGTTCCTCCACTGCTGATGGAGTATGGTGGTG 2543
QY 2217 TCCTCTCTTGTGTCAGACCCCGCGCTCCTCAAGGAGGCCATCCATGTATCAGCTGCG 2276
Db 2544 TTCTCAGTCCGCAACTCCGAGTCTCTTCTGAAAGAGCTATCCATGTTATAAGCTGTG 2603
QY 2277 GCTACGAGACAAGACCGACTGGGGCTGGAGCTGGGGTGGATCTACGGGTGCATCAGCG 2336
Db 2604 GCTATGAGGACAAGACTGAATGGGAACTGAGATCGGGTGGATCTACGGTCTGTGACAG 2663
QY 2337 AGGACATCTGACGGGGTTCAAGATGCACTGCGCGGGTGGCGCTCCGTGTACTGATGTC 2396
Db 2664 AAGACATTTCTACCGGATTCAGATGCAACGGCGAGGCTGGCGGTGATCTACTGATGTC 2723
QY 2397 CGAAGCGGCGGCTTCAAGGGGTCCGCGCGCATCAATCTATCGGACCGTCTCAACGAG 2456
Db 2724 CCAAGCGCGCAGCTTCAAGGGGTCTGCCCATCAATCTTTCGGAACCGTCTGAACCCAGG 2783
QY 2457 TGCTCCGTTGGGCTGGGGTCCGTGAGATCTTCTTACGCGGCAACAGCCCTCTGTGT 2516
Db 2784 TGCTCCGTTGGGCTCTTGGGTCCGTGAGATCTCTTACGCGGCACTGCCCCCTGTGGT 2843
QY 2517 ACGGCTACAAGAACGGCAACCTCAAGTGGCTGGAGCGCTTCGCGCTACATCAACACCA 2576
Db 2844 ACGGCTAC ---GGAGGGCGGCTCAAGTTCCTGAGAGATTCGCGTATCAACACCA 2900
QY 2577 TCTACCCCTTCACTCTCGCTCCGCTCTCGCTTACTGCAACCTTCGCCCGCGTCTGCTCC 2636
Db 2901 TCTACCCGCTCAGTCCATCCCGCTTCTCATCTACTGCACTCTGCGCGCATCTGTCTGC 2960
QY 2637 TCACCGGCAAGTTTCATCATGCTCGATAGCACGTTTCGCGAGCTTCTTCTTCTTCTCC 2696
Db 2961 TCACCGGAAAGTTTCATCTCCAGAGATCAGCAACCTTCGCGCAGCATCTGGTTCTATCTCC 3020
QY 2697 TCTTCATGTCATCTTTCGCGACCGGCACTCTGAGATGCGGTGGAGCGGGTGGAGCATCG 2756
Db 3021 TCTTCATCTCGATCTTCGCCACCGGCACTCTGAGATGAGTGGAGCGGGTGGGATCG 3080
QY 2757 AGGAGTGGTGGAGGAACGACGAGTTCGTGGGTCAATCGCGCGCGTGTCCGCGCATCTCTTCG 2816
Db 3081 ACGAGTGGTGGAGGAACGACGAGTTCGTGGGTGATCGGGGCACTCTCGCGCATCTCTTCG 3140
QY 2817 CGGTCGTGAGGGCTCTCAGGTCTCGCGGATCGACACCAACTTCACCGTCACT 2876

```
Db 3141 CCGTGTTCAGGGCTGTCAAGTGTGTGGCGGCGATCGACCACTTACCGTCACCT 3200
Qy 2877 CCAAGGCCACCGGCGACGAGGACGACGAGTTTCGCGAGCTCTACGCCCTTCAAGTGGACCA 2936
Db 3201 CCAAGGGCTCG---GACGAGGACGGCGACTTCGCGGAGCTGTACATGTTCAAGTGGACGA 3257
Qy 2937 CGCTCCTCATCCCGCCACGACGCTGTCTCATCATTAAGTTCATCGGGGTGCTGGCGCGCA 2996
Db 3258 CGCTCCTGATCCCGCCACCAACCATCCTGATCATCAACCTGGTCGGGGTCTGCGCGGCA 3317
Qy 2997 TCTCCGACGCCATCAACAAACGGGTACCAAGTCTCTGGGGGCCCTCTTCGGCAAGCTCTTCT 3056
Db 3318 TCTCCTAGCCATCAACAGCGGATACCAAGTCTGTGGGGCCCGCTCTTCGGCAAGCTCTTCT 3377
Qy 3057 TCGCCTTCTGGGTTCATCGTCCACCTCTACCCGTTCTCAAGGGGCTCATGGGGCGCCAGA 3116
Db 3378 TCGCCTTCTGGGTTCATCGTCCACCTGTACCCGTTCTCAAGGGGCTCATGGGGCAGGAGA 3437
Qy 3117 ACAGGACGCCACCGTTGTTGTCATCTGTCATCTGCTGGGCCCTCCATCTTCTCCCTGC 3176
Db 3438 ACCGACCCCGACCATCGTGTCTGTCTGGGCCATCCTGCTGGCGTCCATCTTCTCCCTGC 3497
Qy 3177 TCTGGGTTCAGATCGACCCCTTTTCATCGTCAGGACCAAGGGGCCCGGACGTCAGGCGAGTG 3236
Db 3498 TGTGGGTTCGATCGACCCCTTACACCCCGGTCATGGGCCCGGATACCCAGACGTTG 3557
Qy 3237 GCATCAATTGCT 3248
Db 3558 GCATCAACTGCT 3569
```

Search completed: March 12, 2005, 02:52:34
Job time : 1256 secs

This Page Blank (uspto)

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
C	1	2689.8	78.1	3168	9	CL976367
	2	1398.8	40.6	3763	3	CL976367 OsIFCC028
	3	1347.6	39.1	3788	3	AY108113 Zea mays
	4	1272.6	37.0	3192	9	AY103701
	5	1266.4	36.8	3783	3	CL972423 OsIFCC022
	6	1259.4	36.6	3728	3	AY104730 Zea mays
	7	1201.4	34.9	3897	3	AY112236 Zea mays
	8	1197.2	34.8	3898	3	AY103655 Zea mays
	9	1132.8	32.9	3696	3	AY110415 Zea mays
	10	1045.2	30.4	3911	3	AY110079 Zea mays
C	11	1033	30.0	2872	3	CNS0A0ML Arabidops
	12	791.2	23.0	2595	9	AY104236 Zea mays
	13	782.4	22.7	1874	6	CL964957 OsIFCC011
	14	757	22.0	789	7	CD726831 EST027 Cu
	15	747.6	22.1	943	9	CF430024 PH1_25 CO
	16	707.6	20.6	1189	3	CG454544 PH1RA96TD
	17	695.8	20.2	872	8	AY107656 Zea mays
	18	680.4	19.8	834	7	CC427726 PUHAM29TD
	19	679.4	19.7	710	6	COS19653 3530_1_12
	20	679.2	19.7	3474	9	CA9261021 SCQGR7104
C	21	670.8	19.5	949	9	CL978864 OsIFCC044
	22	646.4	18.8	822	9	CG2111396 OGSDF507C
	23	645.2	18.7	687	6	CG198455 PG5RA96TD
	24	644.6	18.7	706	6	CA127439 SCCCLR32CO
						CA155363 SCACRZ310

Db 61 CGGGGACAGGAGCCCAAGCCGCTCGGGCGCTGAGCGGCAGGTGTGCGAGATATGC 120
Qy 212 GGGCAGAGGTCGGGCTCAAGGTGAGCGGGAGCCTCTTCTGCTGCTGCAACAGTTCGGC 271
Db 121 GGGCAGAGGTCGGGCGGCAACCGTCGACGGCGACCTCTTCTGCTGCTGCAACAGTTCGGC 180
Qy 272 TTCCCGGTGTGCGGGCCCTGCTACGAGTACGAGCGCGGGAGGCGACGCAACTGCCCC 331
Db 181 TTCCCGGTGTGCGGGCCCTGCTACGAGTACGAGCGCGGGAGGCGACCCAGAACTGCCCC 240
Qy 332 CAGTGAAGAAGCGGCTACAAGCGCCTCAAGGGAGGCCCGAGGGTTCGCGGGACGATGAC 391
Db 241 CAGTGAAGAAGCGGCTACAAGCGCCTCAAGGGAGGCCCGAGGGTTCGCGGGAGCAGGAC 300
Qy 392 GAGGAGGACATGACAGACCTGGAGCAGGATTCACATCGACGACGAGATCAGCAGAGG 451
Db 301 GAGGAGGACATTCAGACCTGGAGCAGGATTCACATCGACGACGAGATCAGCAGAGG 360
Qy 452 CAGCTGAGGGCAA-----CATGCAGAACAGCCAGATCACCGAGGCGATGCTGCAACGGC 505
Db 361 CTGACAGGATCAGGATGGCATGCGAAGCAGGCAATCACCGAGGGGATGCTGCAACGGC 420
Qy 506 AGGATGAGCTACGGGAGGGGCCCGCAGCGGCGAGCGCAACCAACCCCGCAGATCCCG 565
Db 421 AAGATGAGCTACGGGAGGGGCCCGCAGCGGCGAGCGGCAACCAACCCCGCAGATCCCG 477
Qy 566 CCATCATACCGGCTCCCGCTCCGCTCCGCTGAGCGGTGAGTTCCGATTAACAACGGG 625
Db 478 CCATCATACCGGCTCCCGCTCCGCTGAGCGGGAGTTCCGATTAACAACGGG 537
Qy 626 TATGGCCACGGCAGGTCGCTCTCCCTGCAAGCGCATCCATCCGTACCCGTGTCT 685
Db 538 CATGGCCATGGCAGTTCTCTCTTCTCCCTGCAAGCGCATCCATCCGTACCCGTGTCT 597
Qy 686 GAGCCAGGAGTCCCAAGTGGGAGGAGAAAGAGTGAAGTGGAGGAGAGATGGAC 745
Db 598 GAGCCAGGAGTCCCAAGTGGGAGGAGAAAGAGTGAAGTGGAGGAGAGATGGAC 657
Qy 746 GACTGGAAGTCCCAAGCGGGCATCTCGCGGGCGGG-CCGATCCGGAAGACATGGAC 802
Db 658 GACTGGAAGTCCCAAGCGGGCATCTCGCGGGCGGGCCCGATCCCGACGACTACGAC 717
Qy 803 GCGCAGTGGCATGAACGAGAGGCGAGGAGCGCGCTGTCGAGGAGGTGTCGATCGCG 862
Db 718 GCGCAGTGGCATGAACGAGAGGCGAGGAGCGCGCTGTCGAGGAGGTGTCGATCGCG 777
Qy 863 TCAGCAAGGTGAACCGGTACCGGATGGTGATCGTGGTGGCTCTCGTTGTGCTCGCCTTC 922
Db 778 TCAGCAAGGTGAACCGGTACCGGATGGTGATCATCTCGTCTCGTGGTGGCTCGGCTTC 837
Qy 923 TTCTCCGGTACCGGTATCTGCAACCGTCCCGGAGCGCATCGGGCTGTGGCTCGTCTCC 982
Db 838 TTCTCCGGTACCGGTATCTGCAACCGTCCCGGAGCGCATCCCGCTGTGGCTCACTCC 897
Qy 983 ATCATCTGCGAGATCTGTTTCGCCATCTCTGAGATCTCGACAGTTCGCCAAGTGGTTC 1042
Db 898 ATCATCTGCGAGATCTGTTTCGCCGTGTGAGATCTCGACCGATTCGCCAAGTGGTTC 957
Qy 1043 CCCATCGACCGCGAGCTACCTCGACCGCTCTCCCTCAGGTACGAGAGGGAAGGGGAG 1102
Db 958 CCCATCGACAGGAGACTTACCTCGACCGCTCTCCCTCGCTACGAGCGGAGGGGAG 1017
Qy 1103 CCGTCCGTGTGTCGGCGGTGGACCTGTTGAGGACAGGTGAGACCGCTCAAGGAGCGG 1162
Db 1018 CCGTCCGTGTATCGGCGGTGGACCTGTTGTCAGCAGCGTGGATCCGCTCAAGGAGCGG 1077
Qy 1163 CCGTGGTGAACCGCAACCGCTCTCATCTCCCGGTGAGTACTACCCCGCTGGACAAG 1222
Db 1078 CCGTGGTGAACCGCAACCGCTCTCATCTCCCGGTGAGTACTACCCCGCTGGACAAG 1137
Qy 1223 GTCTCTCTGCTACGCTCCGACGAGCGGCGCTCGATGCTGAGCTGCTGCTGCTCGGAG 1282

Db 1138 GTGTCTCTGCTACTGTTCTCGACGAGCGGCGCTCCATGCTCACGTTGAGTGTGTCGGAG 1197
Qy 1283 AGGCGCGAGTTCGCGCGCAAGTGGGTGCCCTTCTGCAAGAAAGTTCGGCATCGAGCCCCGC 1342
Db 1198 AGGCGCGAGTTCGCGCGCAAGTGGGTGCCATTTCTGCAAGAAAGTTCAGCATCGAGCCCCGC 1257
Qy 1343 GCGCGCGAGTTCCTACTTCTCGCTCAAGGTGCGACTACTCTCAAGGACAAAGGTGCAAGCCACC 1402
Db 1258 GCGCGCGAGTTCCTACTTCTCCAGAAAGTTCGACTACTCTCAAGGACAAAGTTCATCCCAAC 1317
Qy 1403 TTCTGTCAGGAGCGCGCGCCATGAAGAGAGATGAGAGGTTCAGAGTTCAGGTCCGGATCAAC 1462
Db 1318 TTCTGTCAGGAGCGCGCGCCATGAAGAGAGATACGAGGATTCAGGTGAGGATCAAC 1377
Qy 1463 GCGCTGTCGCGCAAGGCCATGAAGGTGCGGAGAGGGGTGGATCATGAAGGACGGCAGC 1522
Db 1378 GCGCTGTCGCGAGGCGCAGAAAGTTCGCGCGGAAGGGTGGATCATGAAGGACGGGAGC 1437
Qy 1523 CCGTGGCCCGGGAAACAACCCCGCAGCAACCCCGGCGATGATCCAGGTGTTCTTGGGCGCAC 1582
Db 1438 CCATGGCCCGGGAAACAACCCCGGCGATGATCCAGGTGTTCTTGGGCGCAC 1497
Qy 1583 AGGCGCGCGCAACGACCGAGGCAACGAGCTGCCCGCGCTCGTGTACGTTCTCCCGTGA 1642
Db 1498 AGGCGCGCGCACGACACCGAGGGCAACGAGCTCCCCCGCTCGTCTACGTTCTCCCGTGA 1557
Qy 1643 AAGCGCCCGGGATTCGAGCACCAAGAAAGCGCGCGCATGAACGCTCTGTGTTCTCGCGTC 1702
Db 1558 AAGCGCCCGGCTTCAGCACCAAGAAAGCGCGCGCATGAACGCTCTGTGTTCTCGTGTG 1617
Qy 1703 TCGCGCTGCTCAACCAACCGCGCTCATCTCAACTTGGACTGTGATCACTACATCAAC 1762
Db 1618 TCGCGCTGCTCAACCAACCGCGCTCATCTCAACTTGGATTTGCGATCACTACATCAAC 1677
Qy 1763 AACAGAAAGCGCATCCGGAAGCGCATGTGTTCTCATGAGACCTCAGTTCGCGCGGAAG 1822
Db 1678 AACAGAAAGCGCATCAGGAGGCGATGTGTTCTCATGAGACCGCGAGTTCGAGCGAAG 1737
Qy 1823 GTCTGTCTGCTCAGTTCCGCGAGGTTTCGCGCATCGACGTGACACCGCATACGCT 1882
Db 1738 GTGTGTCTGCTGAGTTCCGCGAGGTTTCGAGCGCATCGACGTCCAGGCCATACGCC 1797
Qy 1883 AACAGAAACCGCTTCTTCGACATCAACATGAAGGGGTGAGCGGCATCCAAAGGCCG 1942
Db 1798 AACCGCAACACCGTCTTCTTCGACATCAACATGAAGGGGTTCATGGATCCAGGGCGCG 1857
Qy 1943 GTGTAGTTCGGAACCGGTGCGTTCCGCGCGCAGGCGCTCTACGCTACACCCCTCC 2002
Db 1858 GTGTAGTTCGGAACCGGTGCGTTTCAGCGCGCAGGCGCTGATCGGATCAACCCCTCC 1917
Qy 2003 AAGGGACCCAAAGGCGCCAAAGATGGTGACCTGCGACTGCTGCCCGCTGCTTCGGCGCGCAAG 2062
Db 1918 AAGGGACCCAAAGGCGCCAAAGATGGTGACCTGCGACTGCTGCCCGCTTCGTTGGGAGGAG 1977
Qy 2063 AAGCGGAAACACCGCCAAAGGACCGGCTCCCGGAGCGCGTCCCGCGCAGCGCGGAGTGGAC 2037
Db 1978 AAGCGGAGCACCGCAAGGACCGGCTCCCGGAGCGCGTCCCGCGCAGCGCGGAGTGGAC 2037
Qy 2120 AGCGACAGGAGATGCTCATGTCCCATGAATTCGAGAGAGGTTTCGGGCGAGTCCGCG 2179
Db 2038 AGCGACAGGAGATGCTCATGTCCGAGATGAATTCGAGAGAGGTTTCGGGCGAGTCCGCG 2097
Qy 2180 GCGTTCGTCAGCTGACGCTGATGGAGAAAGGGCGCTCCCTCTCTGTCGAGCGCGCGCC 2239
Db 2098 GCGTTCGTCAGCTGACGCTGATGGAGAAAGGGCGCTCCCGCGCTGCTCCAGCGCGCGCC 2157
Qy 2240 GCGCTCTCAAGGAGGCGCATTCATGTGTCATCAGCTGCGGCTTCAGGAGCAAGACCGACTGG 2299
Db 2158 GCGCTCTCAAGGAGGCGCATTCATGTGTCATCAGCTGCGGCTTCAGGAGCAAGACCGACTGG 2217
Qy 2300 GCGCTGAGGAGTGGGTGGATCTACGGGTTCATCGGAGGACATCTTCGAGGGGTTCAAG 2359
Db 2218 GGTCTCGAGTGGGTGGATCTACGGGTTCATCGGAGGACATCTTCAGGAGGACATCTTCAGG 2277

Db 805 AGGAGAAATCCCTGGTCTTCCCTGACCATCATATGATGTCCCAACTGGGAACATTGG 864
Qy 648 CTTCCCTGCACAGCGCATCCATCCGTACCGTGTCTGAGCCAGGAGTGCACAGTGG 707
Db 865 CAAGCGTGTCCATTTC - CCTATGTGAACCAATTCGCCAAATCCGTCAAGGAGTTCCTCTG 923
Qy 708 ACGAGAGAAGAAGTGTGAGTCGAAGGAGAGGATGGACGACTGGAAAGTCCAAGCAGGGCA 767
Db 924 GTAGCATTTGGGATGTTGCCGTGGAAGAGGGGTGTGAGTGGCTGGAATAATGACGAGACA 983
Qy 768 -----TCC 770
Db 984 AGGGAGCAATCCCATGACGAATGGCACAGCAATGTCTCCCTCTGAGGGTCCGGGTGTG 1043
Qy 771 TCGCGCGCGCGGATCCGAGAGACATGAGCCCGAGCTGGCACTGMAAGCAGGCGCA 830
Db 1044 GTGATATGATGATCAACTGATTACAAATGGAAGATGCTTATTGAACGACGAACACT 1103
Qy 831 GGCAGCGCTGTTCGAGGAAGGTGTGATCGCGTCGAGCAAGGTGAACCCCGTACCAGATGG 890
Db 1104 GACAGCTCTATCTAGGAAGTTCACCTTCCTTCCAGGATAAATCCATACAGATGG 1163
Qy 891 TGATCGTGTGCTGTGCTGTGCTCGCTTCCTTCCTTCGAGTACCGGTATCCTGCACCCCG 950
Db 1164 TCATTGTGCTGCGATTGATTGTTCTAAGCATCTTCTTGCACTACCGTATCACAAATCCTG 1223
Qy 951 TCCCGGAGCCATCGGGCTGTGCTGTCTCCATCATCTCGAGATCTGGTTCGCCATCT 1010
Db 1224 TCGGCAATGCATACCCATTATGGCTTCTATCTGTATTATGTGAGATCTGGTTCGTCTT 1283
Qy 1011 CTTGGATCTCTGACCAAGTTCGCCAAGTGTTCGCCATCGACCGGAGACGTACTCGACC 1070
Db 1284 CGTGGATATTGGATCAGTTCCTTAAGTGTTCATCAACCGGAGACGTACTTGATA 1343
Qy 1071 GCCTCTCCCTCAGGTACGAGGGAAGGAGCGGTGCTGTGCTGCGCGGTGGACCTGT 1130
Db 1344 GGCTGGCATTAAGTATGACCGGAAGGTGAGCCATCTCAGTTGGCTGTGTTGACATTT 1403
Qy 1131 TCGTGACAGTGTGACCGCTCAAGAGCGCGCGTGTGACGCCAACACACGTCGTCT 1190
Db 1404 TCGTCAGTACAGTCGACCCCAATGAAGAGCTCTCTTGTCACTGCCCAATACCGTGTAT 1463
Qy 1191 CCATCTTCGCGTAGATACCCCGTGGACAAGTCTCTCTGCTAGCTCTCCGACGAGCGG 1250
Db 1464 CCATCTTGTGTGGATTACCTGTGGATAAGTCTCTCTGCTATGATCTGATGATGGAG 1523
Qy 1251 CGTCGATGTGACGTTTCAGTGTGCTGTGAGAGCGCGGATTCGCGCGCAAGTGGGTGC 1310
Db 1524 CTGCGATGTGACATTTGATGCACTAGCTGAGACTTCAGAGTTTGTAGAAAAATGGGTAC 1583
Qy 1311 CCTTCTCAAGAGTTTCGGCATCGAGCCCGCGCGCGGAGTTCTACTTCTCGCTCAAGG 1370
Db 1584 CATTTGTGAAGAAGTACAAATTTGAACCTTAGAGCTCTCTGAATGGTACTTCTCCAGAAA 1643
Qy 1371 TCGACTACCTCAAGGAAGTTCAGCGCCACCTTCGTGCGAGGCGCGCGCCCATGAAGA 1430
Db 1644 TTGATTAATTGAGGAACAAGTGACCCCTCATTTTGTAAAGACCGCGCGCCCATGAGA 1703
Qy 1431 GAGAGTATGAGGATTCAGGTCGAGTCAACGCGTGTGGTGGCAAGGCCATGAAGTGC 1490
Db 1704 GAGAAATGAAGAAATCAAGATTAGGGTAAATGGCCTTGTGCTAAGGCAACAGAAATTC 1763
Qy 1491 GGCAGAGGGGTGATCATGAGACGGCAGCGCGTGGCCCGGGAACAACACCCGCAAC 1550
Db 1764 CTGAGGAAGGATGGATCATGCAAGATGGCACACCATGGCCAGGAAAACAATACMGGGACC 1823
Qy 1551 ACCCGGCATGATCCAGGTGTTCTTGGGCCACAGCGGGGCCAGACACCGAGGGCAAG 1610
Db 1824 ATCTGGAATGATTCAGGTTTCTTGGTCAACAGTGGTGGCTTGATCTAGGGCAATG 1883
Qy 1611 AGTGCCTCCGCTCGTGTAGTGTCTCCGTGAGAAAGCGCCGGGATTCAGGACCAAGA 1670
Db 1884 AGCTACCCCGTTTGGTCTATGTTTCTCGTGAAGAGCGTCTCTGGATTCAGCATCACAGA 1943

Qy 1671 AGGCGGCGCCATGAACGCTCTGATTCGGGTCTCCGCCGTGTGTGACCAACGCGCCATTCA 1730
Db 1944 AAGCTGGTCCCATGATGCTCTTGTTCGTGTCTCAGCTGTCTACCAATGGACAATACA 2003
Qy 1731 TGTCTCAACTTGGACTGTGATCACTACATCAACAGCAAGGCCATCCGGGAGGCCATGT 1790
Db 2004 TGTGGAATCTTGATTTGATCACTACATTAACAACAGTAAGGCTCTCAGGGAAGCTATGT 2063
Qy 1791 GCTTCTCATGACACCTCAGGTCCGCCGAAGTCTGCTGTACGTTCAAGTTCCTCCGACAGGT 1850
Db 2064 GCTTCTTATGACCCCTAACCTAGGAAGGAGTGTCTGTACGTCCAGTTCCTCCAGAGAT 2123
Qy 1851 TCGACGCGCATCGAGCTGACGACCGGATACGCTAAACAGGAACACCGTCTCTCTCGACATCA 1910
Db 2124 TCGATGECATTTGACAGGAATGATCGATATGCCACAGGAACACCGTGTCTTCGATATTA 2183
Qy 1911 ACATGAAGGGCTGGACGCGCATCCAAAGCCCGGTGTACGTGGGAACAGGTGGTGTTC 1970
Db 2184 ACTTGAGAGGTCTTGATGGCATCCAGGACCAAGTTTATGTCCGAACTGGCTGTGTTTTCA 2243
Qy 1971 GSGCCAGGGCTCTACGGCTACACCTCCCAAGGGAACCAAGA----- 2015
Db 2244 ACCGAACAGCTCTATATGTTATGAGCCCCCAATTAAAGCAGAGAAGGGTGTGTTCTTGT 2303
Qy 2016 -----GGCCCAAGATGGTGCACCTGCGACTGTGCTGCCGTGCTTCGGCGCGCAAGA 2063
Db 2304 CATCACTATGTGGCGGTAGGAGAGGCAAGCAAAATCAAAGAGGGCTCGGACAAGAGA 2363
Qy 2064 AGCGGAAAACGCCAAAGGAC-----GGGCTCCCGGAGG 2096
Db 2364 AGTCGAGAAAGCATGTGGACAGTTCTGTGCCAGTATTCAACTTGAAGATATATAGAGAGG 2423
Qy 2097 GCACCGCTGATATGGGAGTAGTAGGACAAGAGAGATGCTCATGTCCCAATCAACTTCG 2156
Db 2424 GAGTTGAAGCGCGTGAATTTGACGACGAGAAATCACTTCTTATGTCTCAAAATGAGCCTGG 2483
Qy 2157 AGAAGCGGTTTCGGGCGAGTCCGCGCGTTCGTCAAGTTCGACGTGATGAGGAAGGCGGCG 2216
Db 2484 AGNAGAGATTTGGCAGTCCGACAGCTTGTGTCCTCCACTCTGATGGAGTAGTGTGGTG 2543
Qy 2217 TCCCTCTTGTGAGAGCCCGCGCTCTCAAGGAGGCGCATTCATGTATCATGAGTGC 2276
Db 2544 TTCTCAGTCCGCAACTCCGGAGTCTCTTCTGAAAGAAAGCTATCCATGTTTATAAGCTGTG 2603
Qy 2277 GCTACAGGACACAGACCGAGCTGGGGCTGGAGCTGGGGTGGATCTACGGGTCCATACAGG 2336
Db 2604 GCTATAGGACAAGAATGAATGGGGAACCTGAGATCGGGTGGATCTACGGTTCGTGACAG 2663
Qy 2337 AGGACATCTGACGGGTTCAAGATGCACCTGCCCGCGGTGGCGCTCCGTGTACTGCAATGC 2396
Db 2664 AAGACATTTCTACCGGATTCAGATGCAACGCGGAGGCTGGCGGTCCGATCTACTGCATGC 2723
Qy 2397 CGAAGCGGCGCGTTCAGAGGCTCGCGCGCGATCAATCTATCGBACCGTCTCAACAGG 2456
Db 2724 CCAAGCGCGCAGCTTTCAGGGGTCTGCCCCCATCAATCTTTCGGGACCGCTCGAAACAGG 2783
Qy 2457 TGTCTCGGTGGGCGCTGGGGTCCGTGAGATCTTCTTCAGCGCGGACAGCCCTGCTGT 2516
Db 2784 TGTCTCGGTGGGCTCTTGGGTCCGTGAGATCTCTTTCAGCCCGGACCTGCCCTCTGTGGT 2843
Qy 2517 ACCGGCTACAAGAACCGCAACTCAAGTGGCTGGAGGCTTCGGCTCATCAACACCAACCA 2576
Db 2844 ACCGGCTAC---GGAGGGCGGCTCAAGTTCCTGGAGAGATTCGGGTACATCAACACCA 2900
Qy 2577 TCTACCCCTTCACTCGCTCCCGTGTGCTGCCCTATGTGACACCTCCCGCGCGCTGTGCCCTC 2636
Db 2901 TCTACCGCTCAGCTCCATCCCGCTTCTCATCTACTGTCATCTCTGCCCGCCATCTGTCTGC 2960
Qy 2637 TCACCGGCAAGTTCATCATGCGCTCGATTAGCAGTTCGCGAGGCTCTTCTTCAATCGCC 2696
Db 2961 TCACCGGAAAGTTCATCATTCCTCAGAGATCAGCAACTTCGCGAGCATCTGGTTCATCTCCC 3020

Qy	820	CGACGAGGCGAGCGCCCTGTGTCGAGGAAGGTGTCGATCGCGTTCGACGAAGGTGAACCC	879
Db	974	GGATGAAGCTAGCGACCCATGTTCGAGAAAGTCCCTATATCATCAACGCCGAAATTAATCC	1033
Qy	880	GTACCGGATGGTGAATCGTGGTGGTCTCGTTGTGTGCGCTTCTTCCTCGGTAACGGTAT	939
Db	1034	CTACAGGATGATTAATCGTTATCCGGTGGTGGTTTTGGGTTTCTTCTTCCACTACCGAGT	1093
Qy	940	CGTGCAACCGTCCCGGACGCCATCGGGCTGGGCTCGTCTCCATCATCTCGGAGATCTG	999
Db	1094	GATGCAATCCGGGAAAGATGCAATTTGCAATTTGGTCTCATATCTGTAATCTGTGAAATCTG	1153
Qy	1000	GTTTCGGCATCTCTCGATCCTCGACAGTTCCTCCCAAGTGGTTCCCATCATCGACCCGAGAC	1059
Db	1154	GTTTGGATGTCCTGTGATCTTGTGATCAGTTCCCAAGTGGCTTCCAACTCGAGAGAGAC	1213
Qy	1060	GTACTCGACCGCTCTCCCTCAGGPAAGAGGAGGAGCGTCTGCTGTCTGCGC	1119
Db	1214	TTACTCGGACCGTTTGTCACTAAGGTTTGACAAGGAAGTCAACCTCTCAGCTTGTCTCC	1273
Qy	1120	GGTGGACCTGTTCTGTCAGCACCGTGGACCGCTCAAGGAGCGCGCTGCTGAGCGCAA	1179
Db	1274	AATCGACTCTTGTGAGTACGGTGTGATCCCAAGGAACCTCCCTTGGTTCACAGCGAA	1333
Qy	1180	CACGCTGCTCTCCATCCTCGCGGTAGACTACCCCGTGGACAGGTCCTCTGTACTGCTC	1239
Db	1334	CATGTCCTTTCATCTTCTGTGATTAATCGGTTGGAAGTCTCTCTGCTATGTTTC	1393
Qy	1240	CGACGACGGCGCTCGATCTGACGTTCCAGTTCGCTGTCGGAGACGGCGCGAGTTCGGCG	1299
Db	1394	TGATGATGGTGTGCAATGCTTACGTTTGAAGCATTTGTCGAAACATCTGAATTTGCAAA	1453
Qy	1300	CAAGTGGGTGCTTCTGCAAGAGTTCGGCATCGAGCCCGCGCCCGGGAGTTCATCTT	1359
Db	1454	GAAATGGGTTCCTTTCAGCAAAAGTTTAATATCGAGCGCTGCTGCTGAGTGGTACTT	1513
Qy	1360	CTCGCTCAAGGTCGACTTACTCAAGGACAAGGTGCAAGCCACTTCTGTCAGGAGCGCGC	1419
Db	1514	CCAACGAGATAGACTACTGAAAGACAAGGTTGCTGCTTCAATTTGTTAGGAGAGAG	1573
Qy	1420	CGCCATGAAGAGAGATGATGAGGATTCAGGTTCGAGTCAACGCGCTGGTGGCGAAGGC	1479
Db	1574	GGCGATGAAGAGAGATACAGGAAATTCAGGTGAAGGATCAATGCTTGGTTGCAAAAGC	1633
Qy	1480	CATGAAGTCCGGCAGAGGGTGGATCATGAAGNCGCACGCGTGGCCCGGGGACAA	1539
Db	1634	CCAAAAGGTTCTGAGGAGGATGGAACAATGCAAGATGGAAGCCCTCGCTTGGAAACAA	1693
Qy	1540	CACCGCGACACCCGGCATGATCCAGGTGTTCTGGGCCACAGCGCGGCGCACGCAC	1599
Db	1694	CGTACGCGATCATCTCTGGAATGATTCAGGATTAATCTTGGCCAAAGTGGCGTGTGATGT	1753
Qy	1600	CGAGGGCAACGAGTGCCTCGCTGTGACGTCCTCGTGAGAGCGCCCGGGATCCA	1659
Db	1754	GGAAGGAAATGAGTTGCTCGCTGCTTATGTCGAGAGAAAAGAGSCCAGGTTATAA	1813
Qy	1660	GCACCAAGAGAGCGCGGCCCATGAACGCTCTGATTTCCGCTCTCGCGCTGCTGACCAA	1719
Db	1814	CCATCAAGAAAGGCTGGTGGCCATGAATGACCTGGTCCGCTGCTCTGCTGTCTTATCAA	1873
Qy	1720	CGGCGCATTCATGCTCAACTTGGACTGTGATCACTACATCAACAACAGCAAGGCCATCCG	1779
Db	1874	TGCTGCAATCACTATTGNACTTGGACTGTGATCACTACATCAACATAGCAAGGCCATAAA	1933
Qy	1780	GGAGGCCATGTGCTTCTCATGGACCTCAGGTTCGGCGGGAAGGTCTGCTACGTTTCACTT	1839
Db	1934	AGAGGCTATGTGTTTCAATGATGGATCCTTTGGTGGGGAAGAAGTGTGCTATGTACAGTT	1993
Qy	1840	CCGCGAGAGGTTTCAGCGGCATTCGACGTGTCAGCACGATACGCTTAACAGGAAACCGCTCTT	1899
Db	1994	CCCTCAGAGGTTTGTATGGTATTCGACAAAAATGATCGCTATCGCTTAAACAGGAACGTTGTCTT	2053

Db 3131 AG-----GAGTTCTCGGAGCTGTACAGTTCAAGTGGACACCCCTGCTGTATACCCCGGA 3184
 Qy 2955 CCACGCTGCTCATCATTAACGTCATCGCGCTCGTGGCCGGGATCTCCGACGCGCATCAACA 3014
 Db 3185 CCACGCTCTCTGCTGAATTCATCGGGTGGTGGCCGGGATCTCGAACGGATCAACA 3244
 Qy 3015 ACGGTACAGTCTCGGGGGCCCTCTTCGCGCAAGCTCTTCTTCCGCTTCTGGGTCATCG 3074
 Db 3245 ACGGTACAGTCTCGGGGGCCCTCTTCGCGCAAGCTCTTCTTCCGCTTCTGGGTCATCG 3304
 Qy 3075 TCCACCTCTACCCGCTCTCAAGGGGCTCATGGGGCCAGAACAGGACGCCACCGCTG 3134
 Db 3305 TCCACCTGTACCCGTTCTCAAGGTCTGGTGGGAGGAGAACAGGACGCCAGATCG 3364
 Qy 3135 TTGTCATCTGTTCATCTCTGCTGCTCCATCTTCTCCCTGCTCTGCGGTCAAGATCGACC 3194
 Db 3365 TCATGCTGTGTTCATCTCTGCTGCTCCATCTTCTCTGCTCTCTGCGGTCCGCTCGACC 3424
 Qy 3195 CTTTCATCTCAGGACCAAGGCGCGGAGCTCAGGACGTGTGGCATCAATTGC 3247
 Db 3425 GTTCTCTCCCAAGAGACGCGCGCTCTGGAGGAGTGTGGCTGGAATGC 3477

RESULT 4

CL972423 3192 bp DNA linear GSS 21-SEP-2004
 OsIFCC022918 Oryza sativa Express Library Oryza sativa (indica
 cultivar-group) genomic, genomic survey sequence.

CL972423.1 GI:52399375

GSS.

Oryza sativa (indica cultivar-group)

Oryza sativa (indica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehartoideae; Oryzaceae; Oryza.

1. (bases 1 to 3192)

Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,

Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,

Wong, G. K. S., Deng, X. W. and Wang, J.

An analysis of transcriptional regulation of the rice genome and

its comparison to Arabidopsis

Unpublished (2004)

Contact: Chen Chen

Department of Bioinformatic

Beijing Institute of Genomics

Chinese Academy of Sciences, Beijing 101300, China

Tel: 86-10-80481559

Fax: 86-10-80488676

Email: chenchen@genomics.org.cn

Rice genomic sequence.

Class: exon-trapped.

Location/Qualifiers

1..3192

/organism="Oryza sativa (indica cultivar-group)"

/mol_type="genomic DNA"

/db_xref="taxon:39946"

/clone_lib="Oryza sativa Express Library"

/note="Oryza sativa exon trapped genomic sequences"

ORIGIN

Query Match 37.0%; Score 1272.6; DB 9; Length 3192;
 Best Local Similarity 64.7%; Pred. No. 2.7e-241;
 Matches 2069; Conservative 0; Mismatches 1029; Indels 99; Gaps 8;

Qy 147 TGATCCGGGGCACAGAGACCCCAAGCCGCTGCGGGCCCTAGCGGGCAGGTGTGCGAGA 206

2 TGGACGGGACGCGGATCCCGTGAAGTGGGGAGGACACGGGAGCGGGCGGCGTCCAGA 61

Qy 207 TATCGGGGACGAGTGGGCTCAGGTGGACGGGACCTCTCGTCGCTGCAACGAGT 266

Db 62 TCTCGGGGACGCGGTGGGACGAGCGGGAGGGCGAGTGTTCGCCGCTGCGAGCT 121

Qy 267 GCGGCTTCCCGTGTGCGGGCCCTGCTACGAGTACGAGCGCGGAGGAGCGACGCAAACT 326
 Db 122 GCGGCTTCCCGTGTGCGGGCCCTGCTACGAGTACGAGCGCAAGGATGCGACCCAGGCTT 181
 Qy 327 GCGGCTTCCCGTGTGCGGGCCCTGCTACGAGTACGAGCGCGGAGGAGCGCGGAGGAG 386
 Db 182 GCGGCTTCCCGTGTGCGGGCCCTGCTACGAGTACGAGCGCGGAGGAGGAGGAGGAGG 241
 Qy 387 ATGACGAG-----GAGGACATCGACGCTCGAGCACCAGTTCACATCG 431
 Db 242 AAGCGGAGGATCTGATGCTGATGTCAGTGACTACAACTACCTGCATCTGCGAGTG 301
 Qy 432 ACGACGGAATACGACGAGGAGGAGTGGAGGCGAATGCGAAGACAGCA-----GATC 484
 Db 302 CCGACCAAGACAGAAATGCTGATAGGATGCGCAGTTGGCGCATGAATGCTGGGGGTG 361
 Qy 485 ACCGAGGAGTCTGCGACGCGGAGGATGAGCTACGGGAGGGGCGCCGACGACGGCAGCG 544
 Db 362 GTGAGATGTGCGGCGTCCCAAGTATGACAGTGGCGAGATCGGGCTCAAAATCTCGGG 421
 Qy 545 AACAAACCCCGCAGATCCCGCCCATCATCACCGGCT-----CCGCTCCGTCGCGG 596
 Db 422 AGAATCCCTGGAGCTTCCCTGATCATATATGATGTCTCTACCGGAAACATTTGCA 481
 Qy 597 TGACGGTGTGTTCCGATTAACCAACGGGTATGGCCACGGCGAGGTCTGCTTCTCCCTGC 656
 Db 482 AGCGTGTCCATTCCTCTATGTGAACCATTCACCAATCCATCAAGGGAGTTCTCTGCG 541
 Qy 657 ACAAGCGATCCATCCGTACCTGTCTGAGCCAGGAGTGCAGTGGGAGCGAGAGA 716
 Db 542 GCATTGGAAATGTTGCTCGGAAAGAAAGATGTGATGGCTGGAAACTGAAGCAGGACA 601
 Qy 717 AAGAAGTGAGCTGGAAGGAGAGATGACGAGTGGAAAGTCCAA---GCAGGCGCATCTCG 773
 Db 602 GAGCAATCCCATGACCAATGGGACAGCATTTGCCCTTCTGAGGTGCGGGAGTTGGTG 661
 Qy 774 GCGCGCGCGGATCCCGAAGACATGGAACGCGAGTGGGCACTGAAACGACGAGCGAGCG 833
 Db 662 ATATCGATGCATCCACTGATTAACAATATGGAAGATGCTTACTGAATGATGAAACTCG 721
 Qy 834 AGCGCTGTGCGAGAAAGTGTGATCGGTCGAGCAAGGTGAACCCGTCACGGATGTGTA 893
 Db 722 AGCCTCTCTAGAAAGATTTCCCTTCTTCAATCAGAAATAAATCCCTACAGAAATGGT 781
 Qy 894 TCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 953
 Db 782 TTGTTCTGCGTTGTTGTTCTAAGCATTTTCTTCACTACCGTATTTACGATCTGTGTC 841
 Qy 954 CGGACGCAATGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1013
 Db 842 GCAATGCGTATCCGCTCTGGCTTTTATCTGTTATATGATGAGATTTGGTTTGTGCTCT 901
 Qy 1014 GGATCTCGACAGTTCCTCAAGTGTTCCTCAAGTGTTCCTCAAGTGTTCCTCAAGTGTTC 1073
 Db 902 GGATTTGGATCAGTTCCCGAAGTGTTCCTCAATCAACCCGCGGAGGAGGAGGAGGAG 961
 Qy 1074 TCTCCCTCAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1133
 Db 962 TGGCATTTGAGGTATGACAGAGAGGTGAGCCATCTCAGTTGGCTGCTGTTGACATTTTG 1021
 Qy 1134 TGACACGGTGGACCCGCTCAAGGAGCGCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1193
 Db 1022 TCAGTACAGTCGACCCCATGAAGGAACTCTCTTGTACTGCGCAATACCGTGTGTGCTCA 1081
 Qy 1194 TCCTCGCGGTAGATACCCCGTGGACAAAGTCTCTCTGCTAGTCTCTCGACGACGCGCGT 1253
 Db 1082 TTCTGTGTGATTACCCAGTGGACAAAGGTCTCTGCTATGATATCTGATGATGATGCTG 1141
 Qy 1254 CGATGTGAGCTTCGAGTCTGCGGAGACGGCGGAGTTCGCGCGCAAGTGGGTGGCTCT 1313
 Db 1142 CAATGCTGACTTTTGTGATGCAATTGGCTGAGACTTCAGAGTTTGTAGAAAGTGGGTTCAT 1201

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

1 (bases 1 to 3783)
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes

JOURNAL

Unpublished (2002)
2 (bases 1 to 3783)
Coe, E.H.

AUTHORS

Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

JOURNAL

Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES

Location/Qualifiers
1..3783
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635765"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus Library"
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Query Match 36.8%; Score 1266.4; DB 3; Length 3783;
Best Local Similarity 63.9%; Pred. No. 4.7e-240;
Matches 2096; Conservative 0; Mismatches 1056; Indels 126; Gaps 7;

QY 87 GCGGATGGAGGCGAGCGGGCTGGTGGCGGCTCGCGGCTCGCAACCGGAACGAGCTGTGC 146
DB 181 GGGGGATGGAGGCGAGCGGGCTGGTGGCGGCTCGCGGCTCGCAACCGGAACGAGCTGTGC 240
QY 147 TGATCCGGGCGACAGGA-----CCCCAAGCGCTGCGGGCGCTGAGCGGGCAGGTGT 200
DB 241 TCATCCCGCGATGGCGAGCGAGGCGCGAAGCCCATGGACCAGCGGAACGGCCAGGTGT 300
QY 201 GCGAGATATGGCGGCGAGGTGGGCTCAGGTGGACCGGCACTTTCGTGCGCTGCA 260
DB 301 GCCAGATTTGGCGGCGAGCGTGGGGCGCAACCCCGAGCGGAGCCGTTCGTGGCTGCA 360
QY 261 ACAGTTCGGCTTCCCGTGGCGGCTTACGAGTACGAGCGCGGCGGGAGGCGGCGC 320
DB 361 ACAGTTCGGCTTCCCGTGGCGGCTTACGAGTACGAGCGCGGCGGGAGGCGGCGC 420
QY 321 AGAACTCCCGCCAGTGAAGACGCGCTACAAGCGCCCTCAAGGGGAGCGCGAGGGTTGCCG 380
DB 421 AGAACTCCCGCCAGTGAAGACGCGCTTCAAGCGCCCTCAAGGGGAGCGCGGCTGCGCG 480
QY 381 GGGACGATGACGAGGAGCATCGACGACCTGGAGCAGGATTCGAACATCGACGAGCA 440
DB 481 GGGACGAGGAGGAGCGGGCTGACGACCTGGAGACGAGTTCAACTGGAGCGCAACG 540
QY 441 ATCAGCAGGCGAGCTGGAGGCAACATGACGACGAGCGAGT----- 483
DB 541 ACGACTCCCGTACCTCGCGAGTCCATGCTCCACGCCCATGAGCTACGCGGCGGCG 600
QY 484 ----CACCGAGGCGGATCTGCGAGCGAGGATGAGCTACGAGGAGGCGGCGCGAGCGCG 539
DB 601 CCGACCTCGAGCGGCTGCGGCGAGCCATTCCACCCCATGCCCAATGTTCCCTCTCACC 660

QY 540 ACGCAACAACACCCCGCAGATCCCGCCCATCATCACGGCTCCCGTCCGTGCGCGTGA 599
DB 661 ACGACAGATGGTGCATGACATCCCGCCGACACGACGCGCTTGTGCCCTCGTTCGTGG 720
QY 600 GCGGTGAGTTTCCGATTAACCAAGGGTATGGCCACGCGGAGGTCTCGTCTCCCTGACACA 659
DB 721 GTGGGGGGGAGAGGATTCACCTCTCCGTACGCGGATCCCAACCTCTCTGTGACAC 780
QY 660 AGCGCATCCATCCGTACCTGTGTCTGAGCCAGGAGTGCACAGTGGGACGAGAAGAAAG 719
DB 781 CGAGGTCTATGGACCCCTTCCAAGGATCTCGCCGATATGGCTACGGGAGCGTAGCATGA 840
QY 720 AAGTGAGCTGGAAGGAGGATGGAACACTGGAAGTCCAAGCGGGGATCTCTCGCGCGG 779
DB 841 AGAGAGGATGGGAGAGCTGGAAGCAGAAAGCAGGAGGAGTGCACACGAGGAAACGATG 900
QY 780 GCGCCGATCCCGAAGACATGAGCGCCGACGTGGCAGTGAACGAGCGCGGAGGCGCGC 839
DB 901 GCGGCGGAGTATGGTGTGATGTCAGATCTACCACTAATGGATGAAGCTTAGACAGCAT 960
QY 840 TGTGAGGAAGTGTTCGATCCGCTCGAGCAAGGTGAACCCGTACCGGATGGTGTGCTGG 899
DB 961 TGTCCAGAAAGATCCCGCTTCTTCAAGCCAAATCAACCCCTATAGGATGATTATAATA 1020
QY 900 TGGCTCTGTTGTGCTCGCCTTCTTCTCGGTACCGTATCTCTGCACCCCGTCCCGGAGC 959
DB 1021 TTGGCTAGTGGTGTGTTGTTCTTCTTCCATACCGAGTGATGTCGCTGCTGATG 1080
QY 960 CCATCGGCTGTGCTGCTCCATCATCTGCGAGATCTGTTCCGCTATCTCTGGATCC 1019
DB 1081 CATTTGTTTATGGCTCATATCTGTGATCTGTGAAATTTGGTTGCCATGTCTTGGATTC 1140
QY 1020 TCGACCAAGTTCCCAAGTGGTTCCCATCGACCGGAGAGCTACCTCGACCGCTCTCCC 1079
DB 1141 TTGACCAGTTTCAAAGTGGTTTCTTATCGAGAGGGAACCTATCTTGAACGCGCTGATT 1200
QY 1080 TCAGGTACGAGAGGAGGAGCGGCTGCTGCTGTGCGGCTGGAGCTGTTCTGTGAGCA 1139
DB 1201 TAAAGTTTGAAGGAGGAGGATCCCTTCTCAACTCGCCCTGTCGATTTCTTGTGAGTA 1260
QY 1140 CGGTGAGCCCGCTCAAGGAGCGCGCTGTGTGACCGCAACACCGTCTCTCCATCTCTCG 1199
DB 1261 CGGTGATCCCTTGAAGGAACCTTCCATTGCTACTGCTAATACTGTTCTATCTATCTTT 1320
QY 1200 CCGTAGACTACCCCTGGACAAGTCTCTGCTACTAGTCTCCGACGAGCGCGCTGATGC 1259
DB 1321 CGGTGGATTTATCCAGTTGATAAGTTTCACTGCTAGGTTTCTGATGATGGTGTGCCATGC 1380
QY 1260 TGAGTTTCGAGTCTGTCGGAGACGCGCGAGTTCGGCGCGAAGTGGGTGCGCTTCTGCA 1319
DB 1381 TGACATTTGAAGCATTTCTGAAACATCTGAAATTTGCAAGAAATGGGTTCTCTTCTGCA 1440
QY 1320 AGAAGTTCCGATTCGAGCGCCCGCGCGAGTTCTACTTCTCGCTCAAGTTCGACTACC 1379
DB 1441 AAGATATAGCTTTGAGCTCTGCTCCAGTGTGTTACTTCCAAAGAAATAGACTACC 1500
QY 1380 TCAGGACAGAGTTCGAGCCCACTTCTGTGAGGAGCGCGCGCATGAAGAGAGATG 1439
DB 1501 TGAAGACAGAGTGGCGCCAACTTTGTTAGAAACGAGAGCAATGAAGAGAGATG 1560
QY 1440 AGAGTTCAAGGTCGGATCAACGCGCTGTGGCAGAGCCATGAAGTGGCGGCAAGG 1499
DB 1561 AGGAATTCAGGTCAGAAATCAATGCTTGTGTTGTTAAAGCCCAAAAGTTCTCTGAGGAG 1620
QY 1500 GGTGGATCATGAAGGACGCGACCGCTGGCGCGGGAACAACACCGCGACACCCCGCA 1559
DB 1621 GATGGCAATGCAAGATGGAATCCATGGCCCGGAAATAATGTCCTGTGATCATCTGGA 1680
QY 1560 TGATCCAGGTGTTCTCTGGGCGACAGCGCGGCCCAACGACCGGAGGCAACGAGTGC 1619
DB 1681 TGATTCAGGTGTTCTCTTGTGTCNAAGTGGTGCCTATGATGTGGAAGGAAATGAGCTG 1740
QY 1620 GCCTGTTGACGTCTCCCGTGAGAGCGCGCGGGATTCAGACCAACAGAGCGCGCG 1679

Db 1741 GATTGGTTTATGTTTCAAGAGAAACAGCGCCAGGCTACAACCATCAACAAGAGCGTGGTG 1800
Qy 1680 CATGAACCGCTGTGATTCGGGTCTCCGGTGTGACCAACAGCGCCATTCATGCTCAACT 1739
Db 1801 CTATGAATGATTTGGTCCGAGTCTCTGCTGTACTAACTAATGCTCTTATTTGCTGAAC 1860
Qy 1740 TGAAGTGTGATCACTACATCAACAACAGCAAGGCCATCCGGAGGCCATGTGCTTCTCA 1799
Db 1861 TGAATTTGTGATCACTATATCAATAATAGTAAGGCTATAAAGGAAGCAATGTGTTTATGA 1920
Qy 1800 TGAACCTTCAGGTCCGCGGAGGTCTGTGCTTACGTTCCGCGAGAGTTTCAGCGCA 1859
Db 1921 TGAATCTTTGCTTGGAAATAAAGTTTGTCTATGTGCAAGTTTCTCAAGATTGATGGGA 1980
Qy 1860 TGAAGTGTGACGACCATAGCTTAACAGGAACACCGTCTTTCGACATCAACATGAAG 1919
Db 1981 TTAATGCCATGATGATATGCTTAACAGAAATGTTGTCTTTTCGATATCAACATGAAG 2040
Qy 1920 GGTGACGCGATCCAAAGCCCGGTGTACGTGGGACAGGTGCGTGTTCGCGCGCAGG 1979
Db 2041 GTTTGAATGGTATCCAGGGCCCAATTTATGTGGGTACTGATGTGTTTCAGAAAGCAGG 2100
Qy 1980 CGCTCTACGCTCAACACCTCCCA 2004
Db 2101 CATATATGGCTACGATGCTCCCAAAAGAAAGCCACCATCAAGAACTTGCAACTGCT 2160
Qy 2005 2033
Db 2161 GGCACAAAGTGGTCAATTTGCTGTGCTTTGTTGTTAAAGAACAGCAAGAACCA 2220
Qy 2034 GCGACTGTGCTCCCGTCTCGCGCGCAAGAG 2084
Db 2221 AGACCTCTAACCTAAATTTGAAGATAAAGAACTTTTAAGAAAGGAATAAATCAAG 2280
Qy 2085 GGTGCGCGAGGAGCCCGCTGATAT 2129
Db 2281 CCCCTGCATATGCTCTTGGTGAATTTGATGAAGCTGCTCCAGGAGCTGAAATGAAGAGG 2340
Qy 2130 AGATGCTATGTCCTCCCATGAACTTCGAGAAAGGTTTCGGGACAGTCCGGCGCTTCGTCA 2189
Db 2341 CTAGTATTTGAATCAACAGAAAGTTTGAAGAAAGAAATTTGGCCAGTCTTCAAGTTTGTG 2400
Qy 2190 CGTGCAGCTGATGAGAGAGCGCGTCTCTCTGTCGAGCCCGCGCTCTCTCA 2249
Db 2401 CATCCACACTTCTTGAGATTTGGTGAACCTGAAGAGTGCAGTCCAGCTTCTCTCTGA 2460
Qy 2250 AGGAGGCCATCCATGTCATCAGCTGCGGCTACGAGGACAAAGCCGACTGCGGCGCTGGAGC 2309
Db 2461 AGGAAGCTATACATGTCATCAGCTGTGGATATGAAGACAAACAGGCTGGGGAAGATA 2520
Qy 2310 TGGGGTGGATCTAGGGTGCATCAGGAGACATCTGACGGGTTCAAGATGCACTGCC 2369
Db 2521 TTGGTTGGATTTATGGATCAGTCACAGAAATATTTACTTGGGTTTAAAGATGCACTGCC 2580
Qy 2370 GCGGGTGGCTCGTGTACTGATGCGAAGCGGGCGGTTCAGAGGGTCCGCGCGCA 2429
Db 2581 ATGGTTGGCGGTCAATTTACTGATACCTAAACGGGCGGCTTCAAGGTTCCGACCTC 2640
Qy 2430 TCAATCTATCGGACCGGTCTCAACAGAGTGTCTCGGGTGGGCGCTCGGTTCGAGATCT 2489
Db 2641 TCAATCTTTCCGATCGTCTTACCAGGTCTTCCGGTGGGCTCTTGGTTCAATGAATTT 2700
Qy 2490 TCTTCAGCGGACAGCGCCCTGTGTACGGCTACAGAGCGGCAACTCAAGTGGGTGG 2549
Db 2701 TCTTCAGCAACCACTGCGCTCTCTGATGAGGT 2757
Qy 2550 AGGCTTGGCTATACATCAACACCACTTACCCCTTCACTCGCTCCGCTGCTCGCT 2609
Db 2758 AAGGTTTGTGATCATTTAACTCCATCGTATACCTTGGACATCTATCCCGCTCTTGGCT 2817
Qy 2610 ACTGCACCTCCCGCGCTGTGCTCTCTACCGGCAAGTTCAATCATCGCTGATGACCA 2669

Db 2818 ATTGCACATTTGCTGCCATCTGCTTCTGACAGGAAATTTATCACGCCAGAGCTTAACA 2877
Qy 2670 CGTTCCGACGCTCTTTCTTCATCGCCCTCTTTCATGCTCCATCTTGGGAGCGGATCCTGG 2729
Db 2878 ATGTTGCCAGCTGTGGTTCAITGTCATCTTTTCACTGCAITTTTGTAGAGATCCTGG 2937
Qy 2730 AGATCGGTGGAGCGGGGTGAGCATCGAGAGTGTGGAGGAACGAGCATTTCTGGGTCA 2789
Db 2938 AATAGATGGAAGTGGTAGCATCGATGCTGGTGGAGAAACGAGCAGTTTGGGTCA 2997
Qy 2790 TCGGCGCGGTGTCGGGCGATCTTCCGCGTGTGAGAGGCGCTGCTCAAGGTCTCGCG 2849
Db 2998 TTTGAGGCGGTGCTTTCACATCTCTTTGCTGTGTTCCAGGAGCTTCTCAAGGTCTAGCTG 3057
Qy 2850 GGATCGACACCAACTTCCCGTCACCTCCAAGGCCACCGCGACGAGGAGCAGAGTTTCG 2909
Db 3058 GTGTAGACAGAGCTTCACTGTGACATCCAG-----GGCGGAGACGAGGAGTTCT 3111
Qy 2910 CCGAGCTTACGCTTCAAGTGGACCAACGCTCTCTCATCCGCGCACACGCTGCTCATCA 2969
Db 3112 CAGAGCTGTACACATTTCAAAATGAGCAGCCCTTCTGATACCTCCGACAAACCTCTCTCTAC 3171
Qy 2970 TTAAAGTATCGGCGTGTGGCGGATCTCCGACGCCATCAACAGCGGTACCACTCT 3029
Db 3172 TGAATCTTCAATGGAGTGTAGCTGGCATCTCCAATGCGATCAACACGAGATGAATCAT 3231
Qy 3030 GGGGGCCCTCTTCGGCAAGCTCTTCTCGCCCTTCTGGGCTCATGCTCCACTTACCCGT 3089
Db 3232 GGGGGCCCTGTTCGGGAAGCTCTTCTTTCGATTTTGGGTGATCGTCCACTTCTTACCCGT 3291
Qy 3090 TCTCAAGGGGTCTATGGGGCGCCAGAACAGGACGCCACCGTTGTGTGTCATCTGGTCCA 3149
Db 3292 TCTCAAGGGTCTGTGGTGGGAGGAGCAGAACAGGACGCCAACGATTTGTCATCTGGTCCA 3351
Qy 3150 TTCTGCTGCTCCATCTTCTCCCTGCTCTGGGTGAGGATCGACCTTTCATCTGTCAGGA 3209
Db 3352 TCTCTGCTGCTGATCTTCTGCTGCTTGGGTCCGATCGACCCGTTCTTGGCGAAGG 3411
Qy 3210 CCAAGGGCCCGGACGTCAGCGAGTGTGGCATCAATTGC 3247
Db 3412 ATGATGTCCTTGGTGGAGGAGTGTGGTCTGATTTGC 3449

RESULT 6

AY112236
LOCUS AY112236 3728 bp mRNA linear HTC 17-OCT-2002
DEFINITION Zea mays CL1160_1 mRNA sequence.
ACCESSION AY112236
VERSION AY112236.1 GI:21216826

KEYWORDS
SOURCE Zea mays

ORGANISM

Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 3728)

AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.

TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes

JOURNAL Unpublished (2002)

REFERENCE 2 (bases 1 to 3728)

AUTHORS Coe, E.H.

TITLE Direct Submission

JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA

COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.

FEATURES
source

Location/Qualifiers
1. .3728
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:630049"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN

Query Match 36.6%; Score 1259.4; DB 3; Length 3728;
Best Local Similarity 61.9%; Pred. No. 1.2e-238;
Matches 2020; Conservative 0; Mismatches 1098; Indels 147; Gaps 7;
Qy 133 GAACGAGTGTGTGATTCGGGGCCACGAGGACCCCAAGCCGCTGCGGGCGCTGAGCGG 192
Db 225 GAGCTCGCTCCCATGGAGGCGACGCGACGGCTGAAGNNNNNNNNNNNNNNNNNNNN 284
Qy 193 GCAGGTGTGCAGATATCGCGGAGGAGTGGGCTCAGCGTGAAGCGGACCTTTCGT 252
Db 285 NNNNNNGTGCAGATCTCGCGGATGCGGTGGGCACTACGCGGAGGAGGACGCTCTTAC 344
Qy 253 CGCTGCAACGAGTGGCTTCCCGTGTGCGGCTTACAGTACGAGCGCGGGA 312
Db 345 CGCTGCACTGTGCGGGTTCGCGGTGCGGCTTACGAGTACGAGCGCAAGGA 404
Qy 313 GGGCAGCAGAACTGCCCCAGTGCAGACGCGCTACAAGCGCTCAAGGGGAGCGCGAG 372
Db 405 CGGCACACAAGCGTGNNNNAGTGNNNNCAAGTACAAGCGCCACGAGGGAGTCCAGC 464
Qy 373 GGTTCGGGACGATGACGAGGAGGACATCGACGCTGAGGACGAGTTCACATC-- 430
Db 465 GATCCGAGGAGGAGGAGGAGGAGTACTGATGCGCGATGATCTAGCGACTTCAACTACC 524
Qy 431 -----GACACGAGAAATCAGCAGAGGACGCTGAGGGCAACATGACAGACGCGAGAT 483
Db 525 TGCATCTGGCAATGACGCCAAGAGCAGAAATGCTGACAGGATGCGAGCTGGCGCAT 584
Qy 484 CACCGAGGCGAT-----GCTGACGCGAGGATGAGTACGCGGAGGCGCCCGACGACGG 537
Db 585 GAATGCTNNNNNACGCGGATGTTGCCGCGCCCAAGATGACAGTGTGAGATCGGGCT 644
Qy 538 CGACGGCAACACACCGCGAGATCCGCGCCATCATCACGGCTCCCGCTCCGTCGCGGT 597
Db 645 TACCAAGTACGACAGTGTGAGATCCCTCGGGGATACATCCNNNNNNNNNNNNNNNN 704
Qy 598 GAGCGGTGAGTTCCGATTACCAACGGGTATGGCCACGGGAGGTCTGCTTCCCTGCA 657
Db 705 NATTTGGGAGAAATCCCTGTGTCTCCCTATATGAATCATTCATCAATCCGTCAGGGAAT 824
Qy 658 CAACGCGCA-----TCCATCCGTACCCCTGTGTCTGAGCCAGGAGTGGCCAA 702
Db 765 CATTTGGCAGGCGCGCCCATTTCCCTATATGAATCATTCATCAATCCGTCAGGGAAT 824
Qy 703 GTGGGACGAGAAAGAGTGTGAGTGAAGGAGGATGACGATGGAAGTCCAAGCA 762
Db 825 CTCTGTGTAGCTTGGGAATGTTGCTGGAAGAGAGGGTTGATGGCTGGAATAATGAAGCA 884
Qy 763 GGGCA----- 767
Db 885 GGAACAAGGAAACAAATCCATGACGAATGGCAACAGCATTTCTCTCTGAGGGCGGGG 944
Qy 768 --TCTCGGGCGGCGCGGATCCGAAGACATGGACCGCGACGTGGCACTGAACGACGA 825
Db 945 TGTGTGTGATATTGATGATCAACTGATTACCAATGGAAGATGCTTTATTAAACGATGA 1004
Qy 826 GGCAGGAGCGCGCTGTGAGAGAGGTGTGATCGCTCGAGCAAGGTGAACCCGTACCG 885

Db 1005 AACTCGCCAGCCTCTATCTAGGAAAGTTCCACTTCTCTCCAGGATAAATCCATACAG 1064
Qy 886 GATGGTGAATCGTGTGCTGCTCGTTCGCTGCGCTTCTCTCCGCTACCGTATCTCTGCA 945
Db 1065 GATGGTGAATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1124
Qy 946 CCCCGTCCCGAGCGCCATCGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1005
Db 1125 TCTGTGCGTAAATGCATACCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1184
Qy 1006 CATCTCTGATCTCTGACCAAGTTCCTCAAGTGGTTCCTCCATCGACCGCGAGAGTACCT 1065
Db 1185 TCTTCTGATATTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1244
Qy 1066 CGACCGCTCTCTCCCTCAGGTACGAGGAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTG 1125
Db 1245 TGATAGACTCGCATTAAGGTATGACCGGGAAGGTGAGCCATCTCAGTTGGTGTGCTG 1304
Qy 1126 CTTGTTCTGAGCAGCGTGGACCCGCTCAAGAGCGCGCTGCTGCTGCTGCTGCTGCTG 1185
Db 1305 CANNNNNGTCACTACTGTCGACCCCAATGAAGGAGCTCTCTTGTCTCACTGCAATACCG 1364
Qy 1186 GCTCTCCATCTCTCGCGTACTACTACCCGCTGGAAGAAGTCTCTGCTGCTGCTGCTG 1245
Db 1365 GCTATCCATTTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1424
Qy 1246 CGGCGCTCGATGCTGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCG 1305
Db 1425 TGGAGTGTCTATGCTGACATTTGATGACCTAGCTGAGACTTCAAGTTCGTAAGAA 1484
Qy 1306 GGTGCCCTCTCGAAGAAAGTTCGCGATCGAGCCCCCGCGCGCTGCTGCTGCTGCTG 1365
Db 1485 GGTGCCATTTGTTAAGAAAGTACAACATTTGAACCTTAGAGCTCTCTGAAATGTTCT 1544
Qy 1366 CAAGGTGCACTACTCAAGGACAAAGTTCGAGCCCCACCTTCGTCAGGAGCGCCGCG 1425
Db 1545 GAAATTTGATTTACTTGAAGGACAAAGTGCACCTTCAATTTGTTAAAGACCGCGCG 1604
Qy 1426 GAAGAGAGATGATGAGGAGTTCAAGGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 1485
Db 1605 NNN 1664
Qy 1486 GGTGCCCGCAGAGGGTGGATCATGAAGGACGCGACCGCTGCGCGCGGGAACAACACCG 1545
Db 1665 NGTCCCTGAGGAGGATGGATCATGCAAGATGGCACCATGCGCAGGAAACAATACCG 1724
Qy 1546 CGACCAACCCCGGATGATCCAGGTTCCTGCGGCAAGCGCGCGGCGCACGACCGAGG 1605
Db 1725 GGAACCATCTCGAATGATTCAGGTTTCTTGGTTCACAGTGGTGTCTGATGATGAGG 1784
Qy 1606 CAACGAGCTGCCCGCTCGTGTACGCTCCCGTGAAGAGCGCGCGGATTCACAGACCA 1665
Db 1785 TAATGAGTACCCGCTTGTCTATGTTTCTGCTGNNNNNGTCTCTGATTCAGGATCA 1844
Qy 1666 CAAGAGGCGCGGCGCCATGAACGCTCTGATTCGCTGCTGCGCTGCTGCTGCTGCTG 1725
Db 1845 CAAGAAAGCTGGTGGCAATGATCTTGTCCGCGCTCTCAGCTGCTTACCAATGACA 1904
Qy 1726 ATTCACTCACTTGGACTGTGATCACTACATCAACAACGAGGAGGCTCCGAGGAGC 1785
Db 1905 ATACATGTTGAATCTTGAATTTGATGATCACTACATCAACAACAGTAAAGCTCTC 1964
Qy 1786 CATGTGCTCTCATGAGCCCTCAGGTGCGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTG 1845
Db 1965 TATGTGCTCTCTTATGATCTTAACCTAGGAGAGGCTGCTGCTGCTGCTGCTGCTG 2024
Qy 1846 GAGGTTCCGCGGATCGACGCTGACGACCGATAGCTTAACAGGAAACCGCTCTCTTTCGA 1905
Db 2025 GAGGTTCCGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2084
Qy 1906 CATCAACATGAAGGGGCTGAGCGGATCCAAAGGCGCGGTGATGCTCGGAGACAGGGT 1965

2085 TATTAACTTGAGAGGTCTTGATGGCATCCAAGGACAGATTTTATGTGGGCACATGGCTGTGT 2144
QY
1966 GTTCCGGCGCCAGCGCTCTACCGCTACAAACCCCTCCCAAGGGACCCAAAGAGGCCCAAGAT 2025
Db
2145 TTTCAACAGAACAGCTCTATATGTTTATGAGNNNNAAATTAAGCAAAAGAGGGTGGTTT 2204
QY
2026 GGTGACCTGCGACTGCTGCCCGTCTTCGGCCGCAAGAGGCGGAAACACGCCCAAGGACGG 2085
Db
2205 CTTGTGTCATCACTATGTGNN 2264
QY
2086 GTTGCC----- 2091
Db
2265 GNNNTAGA 2324
QY
2092 GGAGGCGACCGCTGATATGGAGTAGATACGCAAGAGATGCTCATGTGCCACATGAA 2151
Db
2325 GGAGGGAGTTGAAGGCGCTGGATTTGATGATGAGAAATCACTTCTTATGTCTCAATATGAG 2384
QY
2152 CTTTCGAGAAAGCGGTTCCGGCAGTCCGGCGGTTTCGTCACCGTCGACGCTGATCGAGGAAGG 2211
Db
2385 CTTTGGAGAGAGATTGGCCAACTCGCAGCTTTTGTGCTCCACTCTCTGATGGAATNNN 2444
QY
2212 CGCGCTCCCTCTTCTGTCGAGCCCGCCGCTCTCCTCAAGAGGCCATCCATGTCATCAG 2271
Db
2445 NNTATC 2504
QY
2272 CTGGGCTACGAGCAAGACCGACTGGGGCTGGAGCTGGGGTGGATCTACGGGTCCGAT 2331
Db
2505 TTGTGGCTACGAGCAAGATTGAATGGGAACTGAGATTGGGTGGATCTATGTTCTGT 2564
QY
2332 CACGGAGGACATCTCTACCGGGTTCAAGATGCACTGCGCGGGTGGCGCTCCGCTGACTG 2391
Db
2565 GACGGAAGATTTCTCACTGGTTCAAGATGCAAGGAGGCTGGCGGTGATCTACTG 2624
QY
2392 CATGCCGAAGCGCGCGCTTCAAGGGTCCGGCGCGATCAATATATGAGACCGTCTCAA 2451
Db
2625 CATGCCCTAAGCGCGCGCTTCAAGGGATCGGCTCCCATCAATCTCTCAGACCGTCTGAA 2684
QY
2452 CCAAGTCTCCGGTGGCGCTGGGTCGTCGAGATCTCTTCAGCGCGGCACAGCCCT 2511
Db
2685 CCAAGTCTCCGGTGGCGCTTCGGTTCAAGTGAATCTTTTACCGCGGATTTGNNNNT 2744
QY
2512 GCTGTACGGCTACAGAACCGCAACCTCAAGTGGCTGGAGCGCTTCGCCCTACATCAACAC 2571
Db
2745 ATGTACGGGTAC--GGAGAGCGCTGAAGTTCTTGAGAGATTCGCCCTACATCAACAC 2801
QY
2572 CACATCTACCTTCACTGCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2631
Db
2802 CACCATCTACCGCTCAGCTCCCTCCCGCTCCTCAATTTACTGTATCTGCTGCCATCTG 2861
QY
2632 CCTCCTCACCGCAAGTTTCATGCGCTCGATTAGCACGTTCCGCCAGCGCTCTTCTTCAT 2691
Db
2862 CCTGCTCACGGGAAGTTTCATCATTCCAGAGATCAGCAACTTCGCTAGTATCTGTTTCAT 2921
QY
2692 CGCCCTCTTCATGTCATCTTCGCGAGCGGCGATCCTTGAGATGCGGTGGAGCGGGGTGAG 2751
Db
2922 CTCTCTCTTCATCTCGATCTTCGCGACCGGTATCTTGAGATGAGGTGGAGCGGCTGGG 2981
QY
2752 CATCGAGAGTGTGGAGGAACAGAGCTTCTGGGTCAATCGCGCGGCTGTCGCGGATCT 2811
Db
2982 CATCGAGAGTGTGGAGGAACAGAGCTTCTGGGTCAATCGCGGCGATCTCCGCCACCT 3041
QY
2812 CTTTCGCGCTGTGCGAGCGCTCTCAAGTCTTCGCGGGATCGACACCACTTCCCGT 2871
Db
3042 CTTTCGCGCTGTTCAGGGCTCTTCAAGGTCTTCGCGGGATCGACACCACTTCCCGT 3101
QY
2872 CACCTCAAGGCCACCGCGGACGAGGACGAGTTCGCGGAGCTCTACGCTCTTCAAGTG 2931
Db
3102 CACCTCAAGGCCCTCG--GATGAAGACGGGACTTCGCGGAGCTGTACATGTTCAAGTG 3158
QY
2932 GACACGCTCTCATTCGCGGCCACACAGCTGCTCATCATTTACGTCATCGCGCTGCTGGC 2991
Db
3159 GACGACACTTCTGATCCGCGGCCACCAACCATCTGATCATCAACCTGTCGCGGTGTTGTC 3218

QY 2992 CGGCATCTCCGACGCGCATCAACACGGGTACCAGTCTCTGGGGCCCTCTTTCGGCAAGCT 3051
Db 3219 CGGCATCTCTAGGCCATCAACAGCGGTACCAGTCTGTTGGGTTCGCTCTTTCGGCAAGCT 3278
QY 3052 CTTTTCGGCTTCTGGGTGATCGTCCACTCTACCGGTTCTTCAAGGGCTCATGGGGCG 3111
Db 3279 CTTTTCGGCTTCTGGGTGATCGTTCACCTGTACCGGTTCTTCAAGGGCTCTCATGGGTG 3338
QY 3112 CAGAACAGAGCGCCACCGTTGTTGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3171
Db 3339 GCAGAACCGGACCCCGACCATCGTGTGTTCTGTTGGGATCTCTGCTGCTGCTGCTGCTGCT 3398
QY 3172 CTTGCTCTGGGTGAGGATCGACCTTTTCATCTGTCAGGACCAAGGGCCCGGACGTCAGGCA 3231
Db 3399 CTTGCTGTTGGTTCGATCGATCCGTTTACCAACCGGCTCACTGGCCGCGATCTCGAAC 3458
QY 3232 GTGTGGCATCAATTGCTGAGCTGTT 3256
Db 3459 GTGTGGCATCAACTGCTAGGAGGT 3483

RESULT 7
AY103655
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 3897)
Coe, E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.

FEATURES
source
1. 3897
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635496"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN
Query Match 34.9%; Score 1201.4; DB 3; Length 3897;
Best Local Similarity 62.7%; Pred. No. 3.5e-227;
Matches 2049; Conservative 0; Mismatches 1101; Indels 117; Gaps 7;
QY 89 GCGATGAGGCCAGCGCGGCTGGTGGCCGCTTCGACACACCGGAGAGCTGGTGTG 148

Db 153 GCGATGGAGCGAGCGCGGGTGGTGGCCGCTCCCAACCGCAACGAGCTCGTGC 212
QY 149 ATCCGGGGCCACGAGAACCC-----AAGCGCTGCGGGCGCTGAGCGGGCAGGTGTC 202
Db 213 ATCCGCGCGACGGCGATCCCGGGCGAAGCGCGCGGAGCAGAACGGGCAAGTGTGC 272
QY 203 GAGATATCGGCGACGAGGTGCGGTTCACGCTGAGGCGGACCTTCTGTCCTGCAAC 262
Db 273 CAGATTGCGGCGACGAGTGGCCCTTGCCCGCGGGACCCCTTGTGGCTGCAAC 332
QY 263 GAGTGGGCTTCCCGTGTGCGCGCTGCTACGAGTACGAGCGCGGAGGCGACGAG 322
Db 333 GAGTGGGCTTCCCGTGTGCGGAGTCTACGAGTACGAGCGCGGAGGCGACGAG 392
QY 323 AACTGCCCCAGTCAAGACGCGCTACAAGCGCTCAAGGGAGCGCGAGGGTTCGCGG 382
Db 393 AACTGCCCCAGTCAAGACTCGATACAAAGCGCTCAAGGGTGCACAGTGTGACCGGT 452
QY 383 GACGATGACGAGGACATCGACGACCTGGAGCAGGTTCAACATGACGACGAGAA 442
Db 453 GACGAGGAGGAGCGCGTTCGATGACCTGGACAACGAGTTCAACTGGGACGCGCATGAC 512
QY 443 CAGCAGAGGCACTGGAGGCA-----ACATGCAAGACAGCCAGATCACGAGCGGATG 496
Db 513 TCGCAGTGTGGCGAGTCCATGCTCTACGGCCACATGAGTACGGCGTGGAGGTGAC 572
QY 497 CTGCACGCGAGGATGAGCTACGGGAGGGGCCCGACGAGCGGCGCAACCAACCCCG 556
Db 573 CTAATGGCGGCCCAAGCTTTCAGCTCAACCCCAATGTTCCACTCTCAACCAACGG 632
QY 557 CAGATCCCGCCCATCATCACCGGCTCCGCTCGGTGCGGTGAGGCGGTTCGAT 616
Db 633 CAAATGGTGGATGATCCACCGGAGCAGCAGCGGCTGTGCTCTTTCATGAGGTGT 692
QY 617 ACCAAGGTTATGGCCACGCGAGTCTCGTCTTCCCTGCAAGCGCATCATCGTAC 676
Db 693 GGGGAAGAGAGATACATCCCTTCCTTATGCGGATCCAGCTTACTGTGCAACCCAG 752
QY 677 CTTGTGTCTGACCCAGGAGTCCAAAGTGGGACGAGAGAAAGAGTGAAGGAG 736
Db 753 TCTATGGACCCATCAAGGATCTGTCATATGGGTATGGTATGTTGCTTGGAGGAA 812
QY 737 AGGATGAGCACTGAAGTCCAGCAGGCGATCTCGGC-----GGCGCGCC 784
Db 813 CGGATGGAGAAATGGAAGCAGAGACAAGAGAGGATGCACAGCGGGAAATGATGGT 872
QY 785 GATCCCGAAGACATGACCGCCAGCTGGCACTGAACGAGCGGAGCGGCGGTGTCG 844
Db 873 GGTGATGATGGTGCATGCTGATCTACCACTAATGGATGAAGCAAGACAAACCTGCC 932
QY 845 AGGAAGGTGTCGATCGCGTGCAGCAAGGTGAACCGGTACCGGATGGTATCGTGGT 904
Db 933 AGGAAATATCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTAATATT 992
QY 905 CTCGTGTGCTCGCTTCTTCTCGGTACCGTATCTGCAACCCGTCGCGACGCCATC 964
Db 993 CTTGTGGTTTTGGGGTCTTCTTCCACTACCGAGTGTGATCGCGGTGAATGATGCA 1052
QY 965 GGGCTGTGCTGCTCTCCATCATCTGCGAGATCTGGTTCGCCATCTCTGGATCTCGAC 1024
Db 1053 GCTTTGTGGCGCATATCTGTATCTGTGAAATCTGGTTGGCATGCTTTGGATCTTGAT 1112
QY 1025 CAGTTTCCCAAGTGGTTCCCAATCGACCGGAGAGGTACTCGACCGGCTCTCCCTCAG 1084
Db 1113 CAATTTCCCAAGTGGTTCCCTATTGAGAGAGACTTACCTTAGACGGGTGCTCAGG 1172
QY 1085 TACGAGAGGAAGGGAGCCGTGCTGTGTCGGCGGTGGAACCTGTTGTGAGACGGT 1144
Db 1173 TTCGACAGGAAGGGCAGCCATCTCAACTTGTCTCCAAATGATTTCTTTGTCAGTACG 1232
QY 1145 GACCGCTCAAGGAGCGCGTGTGACCGGCAACCGGTGCTCTCCATCTCTCGCGTA 1204

Db 1233 GATCCCTTAAAGBACCTCTTTGGTCAACAATACTGTCTTCTATCTATCTCTTCGGTG 1292
QY 1205 GACTACCCCGTGGACAAGTCTCTGTCTACGCTCTCCGACGAGCGCGCTGATGCTGACG 1264
Db 1293 GATTATCTGTGTGATAAGGTTCTTGTCTATGTTTCTGATGATGGTGTGCTCAATGTAACG 1352
QY 1265 TTCGAGTCTGTGTCGGAGACGGCGAGTTCGCGCAAGTGGTCCCTCTCTGCAAGAG 1324
Db 1353 TTTGAAGCATTTATCTGAAACATCTGAATTTGCAAGAAATGGGTTCTTCTGCAACGG 1412
QY 1325 TTCGCGCATGAGCCCGCGCCCGGAGTTCTACTTCTCGCTCAAGGTGCACTACCTCAAG 1384
Db 1413 TACATATTGAACCTCGGCTCCAGAGTGGTACTTCCAAACAGAGATAGACTTGTGAA 1472
QY 1385 GACAAGTGTGACGCCACCTTCTGTGAGAGCGCGCCCATGAAGAGAGATATGAGAG 1444
Db 1473 GACAAGGTGGCAGCAAACTTTGTTAGGAGAGGAGCAATGAAGAGAGATATGAGGAA 1532
QY 1445 TTCAGGTCCGATCAACGCGCTGGTGGCAGAGCCATGAAGGTCCCGCAGAGGGTGG 1504
Db 1533 TTCAGGTGAGAACTAATGCTTGTGTCAAAGCCCAAGAAATGTTCTTGAAGAGAGATG 1592
QY 1505 ATCATGAAGGACGGCAGCGCTGGCCCGGGAACAACACCCCGACACCCCGGCATGATC 1564
Db 1593 ACAATGCAAGATGAAACCCCTGGCTGGAACAATGTTGCTGATCATCTTGGAAATGAT 1652
QY 1565 CAGGTGTTCTGGGCCACAGCGCGGCCACGACACCGAGGCAACAGCTGCCCCCGCTC 1624
Db 1653 CAGGTCTTCTTGGCCAAAGCGGAGGCTTGACTGTGAGGAAATGAATGTCACGATG 1712
QY 1625 GTGTACGTCTCCGTGAGAGGGCGCGGATTCCAGCACACACAGAGCGCGGCGCATG 1684
Db 1713 GTTTATGTTTCTAGAGAAACGACAGGCTATAACCATATAAGAAAGCTGGTCTATG 1772
QY 1685 AACGCTCTGATTCGCGTCTCCCGCTGTGACCAAGCGCCATTCATGCTCAACTGTCAC 1744
Db 1773 AATGATTTGGTCCGAGTCTCTGCTACTACAAATGCTCCATATTGTTAAACTTGAT 1832
QY 1745 TGTGATCATCATCAACAACAGCAAGGCCATTCGCGGAGGCAATGCTTCTCTATGAC 1804
Db 1833 TGTGATCACTACATCAACAACAGCAAGGCTATAAGGAAGCAATGTGTTTTATGATGAC 1892
QY 1805 CTTAGGTGCGCGGAGGCTGCTGCTACGTTCAAGTTCGCGCAGAGGTTTCGACGGCATGAC 1864
Db 1893 CCTTTACTAGAAAGAGGTTTGTCTATGATACAGTTCCCTCAAGATTTGATGGGATGAT 1952
QY 1865 GTGCACGACCGATACGCTAAACAGAAACACCGTCTTCTTCGACATCAACATGAAGGGCTG 1924
Db 1953 CGCATGACCATATGCTAACCGGAATGTTGCTCTTTTGATATCAACATGAAAGTTG 2012
QY 1925 GACGGCATCAAGGCCCGGTGTACGTGCGGACAGGGTGTGTTCCGCGCCAGCGCTC 1984
Db 2013 GATGGTATTACAGGTCCAAATTAATGTTGTTACTGATGTGATTTAGAAAGGCGAGCATTA 2072
QY 1985 TACGCTACACCTTCCCAAGGACCCNAGAGGC-----CAAGATG 2026
Db 2073 TATGGTTATGATGCCCCCAAAACAAAGAACCCATCAAGGACTTGCACATGCTGGCCC 2132
QY 2027 GTGACCTGCGACTGCTGCCCTGTTCGGCCGCAAGAGCGGAACACGCCAAGACGGG 2086
Db 2133 AAGTGGTCTTTTGTGTTGCTGTTGGCAATAGGAAGCAAAAGNAGACTACCAACCC 2192
QY 2087 CTGCGGAG----- 2095
Db 2193 AAAACAGAGAGAAAAAGTTATTATTTTCAAGAAAGAGAAACCAATCCCTGCATAT 2252
QY 2096 -----GGCACCGCTGATATGGAGTAGATAGCGACAAAGAGATGCTCATG 2140
Db 2253 GCTCTGTGTGAAATTTGACGAAGCTGCTCCAGGAGCTGAGAAATGAAGGCGCGGTATTGTA 2312
QY 2141 TCCCAATGAACTTCGAGAACGGTTCGGGCACTCCGCGGCTTCTGTCACGCTCGACGCTG 2200
Db 2313 AATCAACAAAAATTAGAAAAAGAAATTTGGCCAAATCTTCTGTTTGTATACACACTT 2372

Qy 2201 ATGAGGAGGCGCGTCCCTTCCTTCGTCGAGCCCGCGCTCCTCAAGAGGCCATC 2260
| | | | |
Db 2373 CTCGAGAATGTGGAACCTTGAAGAGTGAAGTCTGCTTCTCTTTTGAAGAAGCTATA 2432
| | | | |
Qy 2261 CATGTCATCAGCTGGGCTACGAGGACAGACCACTGGGGCTGGAGCTGGGTGATC 2320
| | | | |
Db 2433 CATGTCATAGTTGTGGTTATGAAGACAAAGACAGACTGGGGAAGAGATTGGCTGATC 2492
| | | | |
Qy 2321 TACGGGTGATCAGGAGGACATCTGACGGGGTTCAAGATGCACCTGCGGGGTGGCG 2380
| | | | |
Db 2493 TATGATCAGTTACAGAAGATATCTAACTGGTTCAAGATGCATTTGTCATGGTTGGCG 2552
| | | | |
Qy 2381 TCCGTGTACTCATGCCGAAGCGGCGGCTTCAAGGGGTGGCGCGGATCAATCTATCG 2440
| | | | |
Db 2553 TCAATTTACTGCATACCTAAACGGGTGTCATTCAAAGGTTCTGCACCTCTGAATCTTCA 2612
| | | | |
Qy 2441 GACGCTCAACAGGAGTCTCGGTGGCGTGGGTCGCGGATCTTTCAGCGG 2500
| | | | |
Db 2613 GATCGTTTCAACAGGTCGTTGGTGGGCTCTGGGTCTATTTGAGATCTTCTTCAGCAAT 2672
| | | | |
Qy 2501 CACAGCCCTCTGTACGGCTACAAGAAACGCAACCTCAAGTGGCTGGAGCGCTGCGC 2560
| | | | |
Db 2673 CATTTGCCCTCTTTGTATGGGT---ATGGTGGCGGTCTGAATTTTGGAAAGATTTTCC 2729
| | | | |
Qy 2561 TACATCAACACCAACATCTACCCCTTCACTCGCTCCCGCTGCTGCGCTACTGCACCCCTC 2620
| | | | |
Db 2730 TACATCAACTCATCGTATCTCTTGGACATCTATTCCTCTTGGCTTACTGTACATTG 2789
| | | | |
Qy 2621 CCGCGCTGCTCCTCACCGCAAGTTTCATCATGCGGTGATAGCAGCTTCCGACG 2680
| | | | |
Db 2790 CTTGCCATCTCTTTATTTGACAGGGAATTTATCTACTCCAGAGCTGAATATGTTGCCAGC 2849
| | | | |
Qy 2681 CTCTTCTTCATCGCCTCTTCATGTCATCTTCGCGAGCGGCATCTTCGAGATCGGTGG 2740
| | | | |
Db 2850 CTGTGTTTCATGTCACITTTTATCTGCATTTTGTCTACGACATCTCTAGAAATGATGG 2909
| | | | |
Qy 2741 AGCGGGGTGAGCATCGAGGAGTGTGGAGAACGAGCAGTTCTGGGTATCGCGGGCGTG 2800
| | | | |
Db 2910 AGTGGTGTGGAATGTAGTACTGTTGGAGGAATGAGCAGTTCTGGGTCAATGGAGGTGTG 2969
| | | | |
Qy 2801 TCGCGGATCTCTTCGCGGTGTCAGGGCTGCTCAAGTTCCTCGCGGATCGACAC 2860
| | | | |
Db 2970 TCCTCACACCTCTTGTGTTTCCAGGGAATCTCTCAAGGTCTATAGTGTGATGATACA 3029
| | | | |
Qy 2861 AACTTTCACCGTCACCTCCAGGCGCACCGCGACGAGGATTCGCGAGCTCTAC 2920
| | | | |
Db 3030 AGCTTCACCGTGACATCAAG-----GGTGGAGATGATGAGGATTTCTAGACTATAT 3083
| | | | |
Qy 2921 GCCTTCAAGTGGACACGCTCCTCATCCCGCCACCAAGCTGCTCATCATTAAGCTCATC 2980
| | | | |
Db 3084 ACATTCAAATGGACTACCTTATTGATACCTCTACCACTTGTCTTATTGAACCTTCATT 3143
| | | | |
Qy 2981 GCGCTGTTGGCGGATCTCCGACGCCATCAACAGGTTACAGTCTCGGGGGCCCCC 3040
| | | | |
Db 3144 GGTGTGGTGTGGCGGTTTCAATGCGATCAATAACGGATATGATCATGTGGGGCCCCC 3203
| | | | |
Qy 3041 TTCGCAAGCTCTTCTTCGCTTCTGGGTCTATCGTCCACCTCTACCCGCTTCTCAAGGG 3100
| | | | |
Db 3204 TTTGGGAAGCTATCTTTGGATTTTGGGTGATTTGTCATCTTTATCCCTTTCTCAAGGT 3263
| | | | |
Qy 3101 CTCATGGGGCGCAAGACAGACGCCACCGTTGTTGTCATCTGGTTCATTTCTGTGCC 3160
| | | | |
Db 3264 TTGGTTGAAGGCAAAACAGGACACCAACGATTTGTCTGCTCTGTTCCATTTCTGTGCT 3323
| | | | |
Qy 3161 TCCATCTTCTCCGCTCTGGGTGAGATCAACCTTTTCATCGTCAGGACCAAGGGCCG 3220
| | | | |
Db 3324 TCAATCTTCTCGCTCTTGGGTTGGATTTGATCTTTCTTGCAGAGGATGATGGTCCG 3383
| | | | |
Qy 3221 GACGTGAGGAGTGTGGCATCAATTGC 3247
| | | | |
Db 3384 CTTCTTGAGGAGTGTGGTTTGATTCG 3410
| | | | |

RESULT 8
AY110415
LOCUS Zea mays CL1166_1 mRNA sequence. 3898 bp mRNA linear HTC 17-OCT-2002
DEFINITION Zea mays CL1166_1 mRNA sequence.
ACCESSION AY110415
VERSION AY110415.1 GI:21214824
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 3898)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 3898)
Coe,E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
Location/Qualifiers
1 .3898
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:630057"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

Query Match 34.8%; Score 1197.2; DB 3; Length 3898;
Best Local Similarity 62.5%; Pred. No. 2.4e-226;
Matches 2026; Conservative 0; Mismatches 1130; Indels 85; Gaps 10;
Qy 89 GCGATGAGGCGCAGCGCGGCTGGTGGCGGCTCGCACCAACCGAAGAGCTGGTGTG 148
| | | | |
Db 186 GAGATGCGGCCCAACAGAGGGATGGTGGCGGCTCGCACCAACCGAAGAGTTCGTCTG 245
| | | | |
Qy 149 ATCGGGGCGCACGAGGACCC-----AAGCCGCTGGCGGCTGAGCGGGCAG 196
| | | | |
Db 246 ATCGGCACACGCGCGCAGCGCGCGCGCTAAGCCACGAAGAGTGCAGATGGCGAG 305
| | | | |
Qy 197 GTGTGCCAGATATCGCGGAGGAGTGGGCTACGGTGGACGCGGACCTCTTCGTGCC 256
| | | | |
Db 306 GTCTGCCAGATTTGCGCGGACTCTGTGGGCGTTTCAGCCACTGTGTGATGTCTTTGTGCC 365
| | | | |
Qy 257 TGCAACAGGTGGCGCTTCCCGTGTGGCGGCTGTCTACGAGTACGAGCGCGGAGGGC 316
| | | | |
Db 366 TGCATGAGTGTGCTTCCCTGTCTGCGCCCTTGTCTATGATGACGAGCGAAGAGGGG 425
| | | | |
Qy 317 ACGCAGAACTGCCCCAGTGCAGACGCGCTTAAAGCGCTTCAAGGGAGCCCGAGGTT 376
| | | | |
Db 426 AACCAATGCTGCCCTCAGTGCAGACTAGATACAAGAGACAGAAAGGTAGCCCTCGATT 485
| | | | |
Qy 377 GCCGGGACGATGACGAGGAGGACATCGACGCTGGAGCAGCTGATTCACATCAGCAC 436
| | | | |

Qy	760	-----GCAGGGCATCTCTCGCGCGCGCGCGATCCCGAA	793
Db	920	GGAACAAGCATTTGCTCCATCAGAAGGCGTGGAGTTGCTGATATTGATTTGATCTTCTACTGAT	979
Qy	794	GACATGACGCGCAGCTGGCACTGAAACGACGAGCGGAGGACGCGCTGTTCGAGGAAGGTG	853
Db	980	TATAACATGGAAGATGCTTACTTGAATGATGAAACTCGGCACCTCTATCTAGANAAGTG	1039
Qy	854	TCGATCCGCTCGAGAAAGGTGAACCCGTTACCGGATGTGTGATCGTGTGCTGCTGTGTG	913
Db	1040	CCAAATCCTTCATCCAGAATAAATCCGTACAGAATGCTCATTTGTCTACGTTTGGCTGTT	1099
Qy	914	CTCGCCTTCTCTCCGTCACGATCCTGACACCGCTCCCGAGCGCATCGGGCTGTGG	973
Db	1100	CTATGCATATTCTTTCGCGTACCGTATCACATCCTGTGAACAATGCATATCCACTGTGG	1159
Qy	974	CTCGTCTCCATCATCTCGAGATCTGTTTGCCCATCTCTGGATCTCTCGACCAAGTTCCCC	1033
Db	1160	CTTTTATCCGTCATATGTGAGATCTGGTTGTTCTGCTGGATTTGGATCAGTTCCCA	1219
Qy	1034	AAGTGGTTCCTCATCGACCGGAGAGTACTCTGACCGCTCTCTCCCTCAGGTACGAGAGG	1093
Db	1220	AAGTGGTCCCCAATCAACCGGTGAACATACCTTGATAGACTGGCTTTAAGTATGACCGA	1279
Qy	1094	GAAGGAGCGCTCGCTGCTGTCGGCGGTGGACCTGTTTCGTGAGCAGGTGGACCCGCTC	1153
Db	1280	GAAGTGAACCATCTCAATTAGTCTCTGTTGATANNNNNGTCAGTCTGTGGATCCAAATG	1339
Qy	1154	AAGGAGCGCGCTGTGTGACCGCCAAACACCGTGTCTTCCATCTCTCGCGGTAGACTACCCC	1213
Db	1340	AAGGAGCTCTCTTGTCACTGCAAAATCTGTGCTTTCCATCTTGTCTGATATCCG	1399
Qy	1214	GTGACAAGGTCTCTGCTGAGTCTCGACGACGCGCGTGTGATGTGACGTTTCGAGTGTG	1273
Db	1400	GTTGACAAGGTATCTTCTATGTTTCCGATGATGGAGCTGCTATGCTGACTTTTGTGCT	1459
Qy	1274	CTGTGCGAGACGGCCGAGTTTCGCGCAAGTGGGTGCCCTTCTGCAAGAGTTTCGGCTC	1333
Db	1460	CTCTCTGAAACTTCAGAGTTTGTAGAAAATGGGTTCCGTTCTGTAGAAGATGCAACATA	1519
Qy	1334	GAGCCCGCGCCCGAGTGTCTACTTCTCGCTCAAAGTTCGACTACTTCAAAGGCAAGGTG	1393
Db	1520	GAGCCTANGCCCCCGGAATGTTACTTTGCTCAGAAATTTGATTACTTTGAAGACAAAGTT	1579
Qy	1394	CAGCCCACTTGTGTGTCAGAGCGCGCCCATGAAGAGAGAGTATGAGGAGTTCAAAGTCT	1453
Db	1580	CAAACTCTATTGTGAAAGAACGCGCGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1639
Qy	1454	CGGATCAACGCTGTGTGGCCACAGGCCATGAGGTGCGGACAGAGGGGTGGATCATGAG	1513
Db	1640	CGTATCAATGGTCTTTGTAGCCAAGGCAACNNNNNGTTCCCGAGGAGGATGGATCATGCAA	1699
Qy	1514	GACGGCACGCGTGGCCCGGGAACAAACACCGCGACCAACCCCGGCATGATCCAGGTGTTT	1573
Db	1700	GATGGTACACTTCTGCGCTGGGAACAAATCTAGGACCATCTCTGMAATGATTCAGGTTTC	1759
Qy	1574	CTGGCCACAGCGCGCGCCACGACACCGAGGGCAACGAGCTGCCCGCCCTCGTGTGAGTCT	1633
Db	1760	CTGGGTACAGTGGAGGGCTTGACGTTTGAAGGCAATGAACTTCTCTGTTGGTTTATGTG	1819
Qy	1634	TCCCGTCAGAAGGCGCCGGATTTCCAGCAACACAGAAGCGCGCGCCATGAACGCTCTG	1693
Db	1820	TCTCGTGNNNNGTCTCGGATTCGAACATCAAGAAGGCTGGTGGCATGAATGACATT	1879
Qy	1694	ATTGCGTCTCCGCGTGTGACCAACGCGCCATTTATGCTCAACTTGGACTGTGATCAC	1753
Db	1880	GTTTCGTATCAGCTGTCTTACTTAATGGGCAATACATGTTTGAATCTTGATTTGACCAC	1939
Qy	1754	TACATCAACAAAGCAAGGCCATCCGGAGGCCAATGTGCTTCTCATFGACCTCAGGTC	1813
Db	1940	TACATCAATATAGCAAGGCTCTTCAGAGAAGTATGTCTTCTTATGACCCCAACCTTA	1999

[illegible]

```
Db 3077 GTGCTTGCTGGTATCGACGAGCTTCACTGTCACTCTTAAGGCCAC---TGACGAAGAA 3133
Qy 2900 GACGAGTTCCGCGAGCTTACGGCTTCAAGTGAACACCGTCTCTATCCCGGCCACACG 2959
Db 3134 GGTGATTTTCCGCGAGCTTACATGTTCAAGTGAACACCGTCTCTGATCCCAACCACT 3193
Qy 2960 CTGCTCATCATTAACGTCATCGGCGTCTGCGCGGCGGATCTCCACGCCATCAACAACGG 3019
Db 3194 ATTTTGATCATCAACCTGTCGCGGTGGTGGCATTTCTTACGCAATCAATACGGT 3253
Qy 3020 TACCAGTCTTGGGGGCCCTCTTTCGCAAGCTCTTCTTCCCTTCTGGGTCTATCGTCCAC 3079
Db 3254 TACCAGTCTGCGGAGCTCTTTTCGGAAGCTCTTCTTGGGTCTGGGTGATGTCAC 3313
Qy 3080 CTCTACCGTCTCTAAGGGGCTCATGGGGCGCGAGAACAGGACGCCACCGTTGTGTC 3139
Db 3314 CTGTACCCCTTCTCTAAGGGCTCATGGGGGAGAGAACCGCACGCCGACCATTTGCTT 3373
Qy 3140 ATCTGGTCCATTCGTGGCTCCATCTTCTCCCTGCTCTGGGTGAGGATCGACCTTTC 3199
Db 3374 GTCTGGGCTATCTCTCTTGGCTCGATCTTCTTCCCTGATGTTGGTTCGTATCGATCATTC 3433
Qy 3200 ATCGTTCAGGACCAAGGCCCGGACGTCAGGACGTTGGCATCAATTGCTGAGCTGTTAT 3259
Db 3434 ACCACCGGCTACTGCTCTGATATCGCGAATGTGGCATCAACTGCTAGGATGAGCTG 3493
Qy 3260 TAAGGTTCAAAATTCGGAGCTTG 3283
Db 3494 AAGATAGTTAAAGAGTGGAACTAG 3517

RESULT 10
CNSOAIML
LOCUS
DEFINITION
  Arabidopsis thaliana Full-length cDNA Complete sequence from clone
  GSLTPGH56ZA06 of Hormone Treated Callus of strain col-0 of
  Arabidopsis thaliana (thale cress).
ACCESSION
  BX832166
VERSION
  BX832166.1 GI:42459101
KEYWORDS
  HTc; GSLT cDNA.
SOURCE
  Arabidopsis thaliana (thale cress)
  Arabidopsis thaliana
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
  1 (bases 1 to 3911)
  Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
  Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
  Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
  Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
  A Combined Approach to Evaluate and Improve Arabidopsis Genome
  Annotation
  Unpublished
  2 (bases 1 to 3911)
JOURNAL
  Genoscope.
REFERENCE
  Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
  BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)
  The sequences are based on single pass reads.
  Life Technologies (a division of Invitrogen) members carried out
  full-length libraries construction : Temple G.
  Genoscope members carried out sequencing and annotation : Castelli
  V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
  Schachter V., Weissenbach J., Salanoubat M.
  URGV INRA : Clepet C., Caboche M.
  Annotation is based on the June 2003 version of the Arabidopsis
  genome released by MIPS (Munich Information center for Protein
  Sequences). 5 prime and 3 prime are assembled with Phrap.
  http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
  length
  http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
```

```
FEATURES
source
Location/Qualifiers
  1..3911
  /organism="Arabidopsis thaliana"
  /mol_type="mRNA"
  /strain="Col-0"
  /db_xref="taxon:3702"
  /clone="GSLTPGH56ZA06"
  /issue_type="Hormone Treated Callus"
  /plasmid="pCMVSPORT_6"
  1..3911
  /gene="At5g64740"

ORIGIN
Query Match 30.4%; Score 1045.2; DB 3; Length 3911;
Best Local Similarity 60.7%; Pred. No. 3e-196;
Matches 1955; Conservative 0; Mismatches 1143; Indels 124; Gaps 10;

Qy 92 ATGAGGCGAGCCCGGGCTGTGGCGGCTGCACACCGGAAACGAGCTGGTGTGATC 151
Db 430 ATGAACACCGGTGGTTCGTTAATCGCCGTTCTCACACAGGAATGATGTTGCTCAT 489
Qy 152 CGGGGCCACGAGGACCCC-----AAGCCGCTCGCGGCGCTGAGCGGCGAGGTGTCGAG 205
Db 490 AATGCCGATGAGAATGCCCGAATAGATCAGTCCAAAGAGCTGAGTGACACAGATGTCAA 549
Qy 206 ATATGCGGACGAGGCTCGGGCTCACGGTGGACGGGACCTCTTTCGCTGCAACGAG 265
Db 550 ATCTGCAGAGATGAGATCGAATTGACTGTTGATGGAGAACCGTTTGTGGCATGTAAACGA 609
Qy 266 TCGGGTTTCCCGTGTGGCGGCCCTGCTACGAGTACGAGCGCCGGGAGGCGACGAGAAC 325
Db 610 TGTGCATTCCTGTGTAGACCTTGTATGATACGAAAGACGAGAGCAATCAAGCT 669
Qy 326 TCGCCCCAGTGCAGACGGCTACAAGCGCTCAAGGGAGCCGAGGTTGCGGGGAC 385
Db 670 TGTCCACAGTGCAAAACCCGTTTCAACGCTCTTAAAGAAAGTCAAGAGTTGAAGTGAT 729
Qy 386 GATGACGAGGAGACATCGACGACCTGGAGCAGAGTTCAACATCGACGACGAGAATCAG 445
Db 730 GAAGAGGAGATGACATTGATGATTTAGACAATGAGTTTGAGTATG----GAAATAATG 784
Qy 446 CAGAGGCGCTGAGGGGCAACATGCAAGACAGCCAGATCACCGAGGCGATGCTGCACGGC 505
Db 785 GGATGGAATTGATAGGTTTCTGAAGAGTATGCAATCTCTCGTGCACACTCCGGTTTCC 844
Qy 506 AGGATGAGCTACCGGAGGGGCCCGGACGCGGCAACCAACACCCCGCAGATCCCG 565
Db 845 CACAATCTGATTTGGATTTCAGCTCCACTGGCTCTAGATTCATTTGCTGACTTACGGCG 904
Qy 566 CCCATCATCACCGGCTCCCGCTCCGTCGCGGTGAGCGGTGAGTTTC-----CGATT 616
Db 905 ACGAGGAGTTGAGATTTCTTCTGATAGACATGCTCTATTGTTCTCTTCACTTGGTG 964
Qy 617 ACCAACGGGTATGGCCACGGCGAGGTTCTGCTTCTTCCCTGACAGCGCATCCATCCGTAC 676
Db 965 GTCATGGCATAGAGTTATCTCTGTTTCTTCTTCTGACCCGCGGTGGTGCACATCCAAGG 1024
Qy 677 CCTGTGCTGAGCCAGGAGTGCACAGTGGGACGAGAGAAAGAGTGAAGTGAAGGAG 736
Db 1025 CCTATGTTACTCAGAAAGATCTTGGCGTTTATGTTATGGAAGTGTGCTTGGANAAGAT 1084
Qy 737 AGGATGACGACTGGAAAGTCCAAGCAG-----GGCATCTCGGGCGC 778
Db 1085 CGGATGAGGAATGGAAGAGAAAGCAGATGAGAAACTTCAGGTTGTTAGGCATGAAGGA 1144
Qy 779 GGGCCCGATCCCGAAGA---CATGGACCGCAGCTGCGACTGAACGACGAGCGGCGGAG 835
Db 1145 GATCCTGATTTGAAGATGGTGTGATGCTGATTTTCCAATGATGAGTGAAGGAGGCGAG 1204
Qy 836 CCGCTGTCGAGGAGGTTGATCGCTCGAGCAAGGTGAACCCGTACCCGATGCTGTGATC 895
Db 1205 CCATTGCTAGGAAGATACCAATCAATTCGAGCAAGATAAATCTTACCGATGTTAATT 1264
```

QY 896 GTGGTGGCTCGTTGTGCTCGCTTCTTCCCTCCGGTACCGTATCTCTGACCCCGTCCG 955
DB 1265 GTGCTACGTCTTGTGATCTTCTGGTCTCTTCTTCTACTACCGTATCTTCCACCCCGTCAA 1324
QY 956 GACGCCATCGGCTGTGGCTCGTCTCCATCATCTGCGAGATCTGGTTGCCATCTCCCTGG 1015
DB 1325 GATGCATATGCTTTGTGGCTTATCTCTGTATATGTGAGATATGGTTTGTCTTTCATGG 1384
QY 1016 ATCTCTGACAGTTTCCCAAGTGGTTCCCATTCGACCGCGAGACGTACTCTGACCGCTTC 1075
DB 1385 GTTCTTGTATCACTTCCCTAAATGGTACCTATCGAGCGAGAAACGTACTTGGACCGACTC 1444
QY 1076 TCCCTCAGTACGAGGGAAGGGAGCGCTGCTGCTGCTCGCGGTGACCTGTTCGTG 1135
DB 1445 TCATTAAAGATATGAGAAAGGGAACCGTCGGGACTATCCCTCTTGGATATATTTGTT 1504
QY 1136 AGCAGCGTGGACCCCGTCAAGGAGCGCGCTGGTGACCGCAACACCGTGTCTCCCATC 1195
DB 1505 AGTACAGTGGATCCATTGAAGAGCCTCGCTTATTAATCTGCAAACTACTGTCTTGTCTATT 1564
QY 1196 CTCGCCGTAGACTACCCCGTGAACAAAGTCTCTCTGTACGTCTCCGACGACGCGCGTGC 1255
DB 1565 CTTGTCTGTGATTATCTCTCGATAAAGTGTCTTGTACGTATCTGATGATGGTCTGCT 1624
QY 1256 ATGCTGACGTTTCGAGTCTGTCGGAGACGCGCGAGTTCCGCGCAAGTGGTGGCCCTTC 1315
DB 1625 ATGCTTACTTTCCGAAGCTCTTTCTGAGACCGCTGAATTCGCAAGGAAATGGGTTCTTTTC 1684
QY 1316 TCGAAGAAGTTCGGCATCGAGCCCGCGCCCGGAGTTCTACTCTCGCTCAAGTCTGCAC 1375
DB 1685 TCGAAGAATAATTGTATTGAGCTCTGTCTCCGAATGGTATTCTGCCATNAATGGAC 1744
QY 1376 TACCTCAAGGACAAGTTCAGCCCACTTTCGTGACGAGCGCGGCCCATGAAGAGAGAG 1435
DB 1745 TACTTTGAAGATAAAGTTTCAATCCCGCATTTGTAGGAGCGCGAGCCATGAAGAGAGAT 1804
QY 1436 TATGAGGAGTTCAAGGTCGGATCAACGGCTGTGTGGCCAGGCTCATGAAGTCCGGCA 1495
DB 1805 TATGAAGAAATTCAAAGTAAGATCAATGCTTTAGTAGCAACAGACAGAAAATGCTCTGAA 1864
QY 1496 GAGGGGTGGATCATGAAGGACGGCACCGCTGGCCCGGGAACAACACCCGCGACCAACCC 1555
DB 1865 GATGGGTGGACTATCGAAGACGGTACACTTGGCCCGGTATATAGTGTGGAGATCATCT 1924
QY 1556 GGCATGATCCAGGT- GTTCTCGGGCCACAGCGCGCGCCACGACACCGAGGGCAACGAGCT 1614
DB 1925 GGCATGAATCAGGTCTTTCTTGGAAAGTGACCGTGTCTGTATGTCGAAACAAACGAGTT 1984
QY 1615 GCCCGCCTCGTGTAGCTTCCCGTGAGAGCGCCGGGATTCAGACACCAAGAGGC 1674
DB 1985 GCCTCGATTAGTTTACGTTTCTCGTGAGAGAGACCCGGATTTGATCACCATAGAGGC 2044
QY 1675 CGCGCCATGAACGCTCTGATTTCGCTCTCGCCGTGTGACCAACGCGCATTCATGCT 1734
DB 2045 TGGAGCTATGAATTCCTGTATACGAGTCTCTGGGGTCTTATAAATGCTCTCTTACTTCT 2104
QY 1735 CAACTTGGACTGTGATCACTACATCAACAAAGCAAGGCGCATCCGGGAGGCGCATGTGCTT 1794
DB 2105 GAATGTGATGTGATCACTACATCAACATAGCAAAAGCTCTTAGAGAAGCAATGTGTTT 2164
QY 1795 CCTCATGACCTCATAGTTCGGCGGGAAGGTCTGTACGTTTCAGTTTCCGCGAGAGTTTGA 1854
DB 2165 CATGATGATATCCTCATGTAGGAAAGAAATCTGTTTATGTTTCAGTTTCCCTCAAAAGTTTGA 2224
QY 1855 CGGCATCGACGTGACGACCGATAGCTAAGAGGAACCGTCTTCTTCGATCACTCAACAT 1914
DB 2225 TGGATTGATAGGACAGATCGATCTCAAAATCGCAATGTGTGTTCTTTGATATCAATAT 2284
QY 1915 GAAAGGGGTGGACGGCATCCAAAGGCCCGGTGTACGTTCGGGACAGAGGTGCGTTTCCGGCG 1974
DB 2285 GAAAGGTTTGGATGGGTACAGGGCCTATATACGTTCGGTACAGGTGTGTGTTTTCAGGAG 2344
QY 1975 CCAGGCGCTCTACGGCTACAACCCCTCCCAAGGGACCCCAAGAGGCCCAAGAGTGGTGACCTG 2034

DB 2345 GCAAGCGCTTTACGGATTTGATGCACCGAAGAAGAGGGCCCAACGTAAGACATGCAA 2404
QY 2035 CGACTGTCTCCCGTCTTCGGCC----- 2057
DB 2405 TTGCTGGCCAAATAGGTGTCTCTATGTGTTTGGTTTCAAGAAGAATCGTAAGCAAGAC 2464
QY 2058 -----GCAAGAAGCGAAACACGCAAGGA 2082
DB 2465 AGTGGCTCGGATAAAGAAGAAGAAATAGGAAACGCTCAAAGCAGATCCACGCAATTAGA 2524
QY 2083 CGGGCTCCGGAGGACCGCTGATATGGGAGTAGATAGCAGCAAGAGAGATGCTCATGTC 2142
DB 2525 AAATATCGAAGAGGCGCGCTCACTAAAGGTTCTAAACGTAGAACAGTCAACGAGGCAAT 2584
QY 2143 CCATATGAATCTCGAGAACGCTTCGGGACGTCCGCGCGCTTCGTCAGCTCGACGCTCAT 2202
DB 2585 GCAATGAAGTTGGAGAAGAAATTTGGCAGTCTCTCTGTAATTTGTTGATCTGCGCGTAT 2644
QY 2203 GGAGGAAGCGCGCTCTCTCTCGTTCGAGCCCGCGCGCTCTCAAGAGAGGCCATCA 2262
DB 2645 GGAGAAATGGTGGGATGGCTAGAAACGCAAGCCCGCTTGTCTGCTTAAAGAAAGCCATCCA 2704
QY 2263 TGTCTACGCTCGGCTACGAGGACAGACCGACTGGGGGCTGGAGCTGGGGTGGATCTA 2322
DB 2705 AGTCATTAGTTGCGGATATGAAGATAAACTGAAATGGGGAAAGAGATTTGGGTGGATCTA 2764
QY 2323 CGGGTTCGATCACGAGGACATCTCTGAACGGGTTCGAAGATGCATTCGCCCGGTGGCGTCT 2382
DB 2765 TGGTCTCTTACCGAGATATCTTACGGGTTTAAAGATGATCTCTCATGTTGGAGATC 2824
QY 2383 CGTGTACTGTACGCGAGCGGGCGGTTCAGGGGTTCGCGCGGATCAATCTATCGGA 2442
DB 2825 TGTATTGTATACAAAGTTAGCGCTTTCAAAGGATCAGCTCCAACTCAATCTTTTCGA 2884
QY 2443 CCGTCTCAACAGGTGCTCCGGTGGCGCTGGGGTCCGTTCAGATCTCTTTCAGCCGGA 2502
DB 2885 TCGTCTCCATCAAGTTCTTCGATGGGCGCTTGGGTTCGGATTTTCTTGGAGTAGGA 2944
QY 2503 CAGCCCCCTGTGTACGGCTACAAGAACGGCAACCTCAAGTGGCTGGAGCGCTTCGCCCTA 2562
DB 2945 TTGTCTTATTTGGTATGGT----ATGGAGGTGGGTGAAATGGCTTGAACGGTGTCTCTA 3001
QY 2563 CATCAACACCACTATACCCCTTCACTCGCTCCCGTGTCTGCTACTGACACCTCC 2622
DB 3002 CATTAACCTCTGTGTTTACCGTGCACCTCTCTACCGCTCATCGTTACTGTCTCTCTCCC 3061
QY 2623 CGCGTCTGCTCTCTCACCGCAAGTTTCATCATGCGCTCGATTTAGCAGCTTTCGCCAGCT 2682
DB 3062 TGCCATCTGTCTTCTCACTGGAAAATTCATCGTTTCCGAGATTAGCAACTATGCGAGTAT 3121
QY 2683 CTTCTTCATCGCCCTCTTTCATGTCCATCTTCGCGACGGCATCTTGGAGATGCGGTGGAG 2742
DB 3122 CCTCTTATGGCGCTCTTCTCGTGCATTTGCAATTAACGGGTATTTCTGAGATGCAATGGG 3181
QY 2743 CGGGGTGAGCATCGAGGAGTGGTGGAGAAACAGCAGTTCT- GGGTCAATCGCGCGCGTGT 2801
DB 3182 CAAAGTTGGGATCGATGATTGGTGGAGAAACGAACAGTTTTTGGGGTCAATTGAGAGTGT 3241
QY 2802 CCGCGCATCTCTTCGCGCTCGTGAGGGCTGTCTCAAGGTCTCTCGCGGGATTCGACCA 2861
DB 3242 CTGCGCATCTGTGTTGCTCTCTTCCAAAGGTCTCTCAAGGTTCCTTGTGTTGTCGACACTA 3301
QY 2862 ACTTCACGCTCACTCCAAAGGCCACCGCGACGAGGACGAGTTCGCCCGAGCTCTACG 2921
DB 3302 ACTTCAGATCAATCAAAAGC-----AGCTGATGATGGAGAGTTCTCTGACCTTTTACC 3355
QY 2922 CTTTCAAGTGGACACGCTCTCTATCCCGCCACACGCTGCTCTCATTAACGTTCATCG 2981
DB 3356 TCTTCAATGGACTTCACTTCTCATCCCTCCCAATGACTCTTACTCATCATAAACGTCATTG 3415
QY 2982 GCGTGTGGCGCGCATCTTCGAGCCCATCAACAAACGGGTACCGATCCAGTCTCTGGGGGCCCTCT 3041

Db 3416 GAGTCATAGTCGGAGTCTCTGATGCCATCAGCAATGGATAGCACTCGTGGGAGCCGGCTTT 3475
Qy
3042 TCGGCAAGCTCTCTTCGCCCTTCCTGGGTCATCGTCCACCTCTACCCGTTCTCAAGGGGC 3101
Db
3476 TCGGAAGACTGTTCTTTGGCACTTTGGGTCATCATCTTTTACCCGTTCTTAAAGTT 3535
Qy
3102 TCATGGGGCGCCAGAACAGGACGCCACCGTTGTTGTATCATCTGTCGATTCATCTGTCGCCCT 3161
Db
3536 TGCCTGGGAACAAGATAGATGCCAACCATTAATTGTCGTCTCCATCCCTCGSCCT 3595
Qy
3162 CCATCTTCTCCTCTCTCGGTCAGGATCGACCTTTTCATCG 3203
Db
3596 CGACTCTTACACTTCTTTGGGCCCGGGTTAATCCGTTTGTGG 3637

RESULT 11

AY104236 2872 bp mRNA linear HTC 16-OCT-2002
LOCUS
DEFINITION Zea mays PC0121439 mRNA sequence.
ACCESSION AY104236
VERSION AY104236.1 GI:21207314
KEYWORDS HTC.
SOURCE Zea mays

ORGANISM

Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS Hayney, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes

JOURNAL

AUTHORS Unpublished (2002)

REFERENCE

AUTHORS 2 (bases 1 to 2872)

TITLE

Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSU, maizeimap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.

FEATURES

source

1..2872
Location/Qualifiers
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:637064"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN

Query Match 30.0%; Score 1033; DB 3; Length 2872;
Best Local Similarity 65.5%; Pred. No. 7,6e-194;
Matches 1634; Conservative 1; Mismatches 796; Indels 64; Gaps 6;
Qy 820 CGACGAGCGAGCGCGCTGTCGAGGAAGGTGTCGATCGGTGCGAGCAAGGTGAACCC 879
Db 6 GGTCCGTGCACGGCTACCTCTAAGTCGCATAGTTCGGATATCTCCAAACGAGCTTAACCT 65
Qy 880 GTACCGATGTGATCGTGGTGGCTCTGTTGCTCGCCCTCTCTCCGTTACCGTAT 939
Db 66 TTATCGGATCGTGATTTGTTCTCCGGCTTATCATCTATGTTCTTCTTCAATATCGTAT 125

Qy 940 CTTGCAACCCGCTCCCGAGCCATCGGGCTGTGGCTCTCTCATCATCTGCGAGATCTG 999
Db
126 AACTCATCCAGTGGGAAGTCTTATGGGTTGTGGCTGTATCTGTATTATTTGTAAGTTG 185
Qy 1000 GTTTCGCCATCTCTCGATCTCGACCAAGTTCCCAAGTGGTTCGCCATTCACCGCGAGAC 1059
Db
186 GTTTCGCTTGTCTTGGCTCTAGATCAGTTCCCAAGTGGTATCTCTATCAACCGTGAAC 245
Qy 1060 GTACCTCGACCGCTCTCCCTCAGGTACGAGAGGGAAGGGAGCCGTCGCTGTCT-GTTCGG 1118
Db
246 TTACCTCGATAGACTTGCATTGAGATATGATGGAGGGTGAAGCATCCAGTTGGGCTC 305
Qy 1119 CGGTGGACCTGTTTCGTAGCACCGTGGACCCGCTCAAGAGAGCCGCGCTGGTGAACCCCA 1178
Db
306 CAATCGATGCTCTTTGTTAGTACAGTGCATCCATTAAAGGAACCTCTCTAATTACTGGCA 365
Qy 1179 ACACGCTGCTCTCATCTCCGCTAGACTACCCGTTGGACAAGGTTCTCTGCTACTCTCT 1238
Db
366 ACATGTCCTGTCCTTCTTGTGCTGGATTTACCTGTGTGACAAAGTATCATGTATTGTTT 425
Qy 1239 CCGACGACGCGCGTCTGATGCTGACGTTTCAGTTCGCTGTCCGAGACGCGCCGAGTTCCGCG 1298
Db
426 CTGATGACGGTTCAGCTATGTTGACTTTTGAAGGGCTATCTGAAAACGCGAGAGTTGCAA 485
Qy 1299 GCAAGTGGGTGCCCTTTTCGAAGAAGTTTCGGCATTCGAGCCCCCGCCCCGGAGTTCTACT 1358
Db
486 GGAATGGGTTCCCTTTTTCGAAGAACACACAATATTGAACCTAGGGCTCCAGAGTTTACT 545
Qy 1359 TCTCGCTCAAGTTCGACTACTCTAAGACAAGGTGCGAGCCACCTTCGTGTCAGGAGCGCC 1418
Db
546 TTGCTCGAAAGATAGATTACCTAAAGGACAAAATACAACTTCTTTTGTGAAGAAGAGGC 605
Qy 1419 GCGCCATGAAGAGAGAGTATGAGAGTTCAAGGTTCGGATCAACGCGCTCGTGGCCCAAGG 1478
Db
606 GGGCTATGAAGAGGAGTGTGAAGAGTTCAAGATCGGATCGATGCTCTGTTGCAAAAG 665
Qy 1479 CCATGAAGTGCCTCGGACAGAGGGTGGATCATGAAGAGCGGACGCGCTGCGCCGGGAACA 1538
Db
666 CGCAAAAATATACCTGAGGAGGGCTGGACCATGGCTGATGCGACTCCCTGCGCTGGGAATA 725
Qy 1539 ACACCCGCGACACCCCGGATGATCCAGGTGTTCTGGGGCCACAGCGGGCCACGACCA 1598
Db
726 ACCCTAGAGATCATTCMGGAATGATCAAGATATTCCTTGGGCGACAGTGGTGGGCTTGACA 785
Qy 1599 CCGAGGGCAACGAGCTGCCCGCGCTCGTGTACGCTCTCCCGTGAAGAGCGCCCGGGATTC 1658
Db
786 CGGATGGGATGAGTTGCCACGGCTTGTATTGTTCTCGTGAAGAGAGCCAGGCTTCC 845
Qy 1659 AGCACCAAGAGAGCGCGCCCATGAACCGCTCTGATTCGGCTCTCCCGCTGCTGACCA 1718
Db
846 AGCACCAAGAGAGGCTGGTGCCATGAATGCTTTGATTGCTGATCAGCTGTCTCTGACGA 905
Qy 1719 ACGGCCATTCATGCTCAACTTGGACTGTGATCACTACATCAACAAGAGAGCCATCC 1778
Db
906 ATGGTGTCTTATCTTCTTAATGTGGATGTGATCACTACTTCAATAGCAGCAAAAGCTCTTA 965
Qy 1779 GGGAGGCGATGTGCTTCTCTCATGGACCTCAGGTTCGCGCGGAAAGGTCTGCTACGTTCAGT 1838
Db
966 GAGAGGCTATGTGTTTCATGATGATCCAGCATAGGAAGGAATACTGCTATGTTCAGT 1025
Qy 1839 TCCCGCAGAGGTTTCGACGGCATCGACGTGCAACCGATACGCTTAAACAGGAACCCGCT 1898
Db
1026 TTCCACAAAGATTGTGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1085
Qy 1899 TCTTCGACATCAACATGAAGGGCTGACCGCATCCAGGCCCGCGGTGATCGTCCGGGACAG 1958
Db
1086 TCTTTGATATTAATATGAAGGGCTTAGATGGCATTTCAAGGACCTGTTTATGTGGGAACAG 1145
Qy 1959 GGTGCTGTTTCCGGCGCCAGCGCTCTACCGGCTACAAACCCCT----- 1999
Db
1146 GATGCTGTTTCAATAGGACGCGCTTGTATGGCTATGATCTCTGTTATGACAGAGCTGATT 1205
Qy 2000 -----CCCAAGGGACCCAAAGAGGCCCAAGATGGTGACCTGCGGACTGCTGCCCCGTG 2049

Db 1206 TGGAGCTTAACATTATTAATAAGTTGCTGTGGCGGAAGAAAAAGAGAGCAAGAGCT 1265
QY 2050 CTTTCGGCCCAAGAACGCGAACAACGCCAAGACCGGCTCGCGAGGGGACCGCTGATAT 2109
Db 1266 ATATTGATTCCAAAGAACCGTGATATGAAGAGAACAGAAATCTTCGGCTCCCATCTTCAACA 1325
QY 2110 GGGAGTAGATAG-----CGAAGAGAGATGCTCATGTGCC 2144
Db 1326 TGAAGATATAGAAGAGGATTTGAAGGTTACGAGGATGAAGAGTCACTGCTTATGTCTC 1385
QY 2145 ACATGAACATTCCAGAAAGCGTTTCGGGACGTCCGCGCGTTCGTACGCTCGACGCTGATGG 2204
Db 1386 AGAAGAGCTTGAGAGAACCGTTTGGCGAGTCTCCAAATTTTATGCAATCCACCTTATAGA 1445
QY 2205 AGAAGGCGGCTCCCTCTCGTAGCCCGCGCGCTCTCAAGGAGGCGCATCCATG 2264
Db 1446 CTCAGAGTGGCATACCCCTTCAACAAACCCAGGTTCCCTGCTAAAGGAAGCTATACATG 1505
QY 2265 TCATCAGCTGGCGCTACGAGGACAAGACCGACTGGGGCTGGAGCTGGGGTGGATCTAG 2324
Db 1506 TCATTAGTTGTGGATATGAGGATAAACAAGAAATGGGGAAAGAGATCGGATGGATATATG 1565
QY 2325 GGTGATCAGCGAGGACATCTGACGGGTTCAAGATGCACTGCCGCGGGTGGCGCTCG 2384
Db 1566 GCTCTGTTACTGAAGATATTTTAATCTGGTTTCAAGATGATGCAAGAGGTTGGATCCA 1625
QY 2385 TGTACTGCATCGAAGCGGGCGGTTCAAGGGGTTCGCGCGGATCAATCTATCGGACC 2444
Db 1626 TCTACTGCATGCCACTTCGGCTTCTTCAAGGGTTCGCTCCGATTAATCTTCTGATC 1685
QY 2445 GTCTCAACAGGTGCTCGGTGGGCGTGGGCTCGTGGAGATCTTTCAGCGCGGACA 2504
Db 1686 GTCTCAACAGGTGATGCGTGGGCTTGTGTTCAAGTTGAATTTACTTAGACACACT 1745
QY 2505 GCGCCCTGCTGACGCTACAGAACGCAACTCAAGTGGCTGAGCGCTTCGCTACA 2564
Db 1746 GTCCTATCTGGTATGGTTAC---AATGAGAGCTAAAGCTTCGAGAGACTGGCATACA 1802
QY 2565 TCAACACCAACCATACCCCTTACCTCGCTCCCGCTGCTCGCTTACTGCACCCCTCCCG 2624
Db 1803 TCAACACCAATTGTTTATCCAAATTACATATCCCACTAGTAGCATACTCGCTCTCTG 1862
QY 2625 CGGTCTGCTCTTCAACGGCAAGTTTCATGCGCTCGCTGCTGCTGCTTACTGCGCAGCTCT 2684
Db 1863 CTATCTGTTTACTCAACCAAAATTTATTTCTCTCGGATAGCAATATGCTGGGGGT 1922
QY 2685 TCTTCATCGCCTCTTTCATGCTCCATCTTCGACCGGATCTGAGATGGGTGGAGG 2744
Db 1923 TCTTCATCTGCTTTTGTCTTCCATCTTCGACCTGGTATTTTGGAGCTTCGATGGAGTG 1982
QY 2745 GGGTGAGCATCGAGGAGTGGTGGAGAACGAGCAGTTCTGGGTCAATCGCGCGGTGTCG 2804
Db 1983 GTGTTGGCATTGAGGATGGTGGAGAAATGAGCAGTTTGGGTCAATGGTGCCACTCTG 2042
QY 2805 CGCATCTCTTCGCGCTGTGCGAGGCGTCTCAAGGCTCTCGCGGATTCGACACCAACT 2864
Db 2043 CACATCTCTTGTCTGTGTTTCAAGGCTCTTAAAGTGTCTAGCAGGATTCGACACAAACT 2102
QY 2865 TCACGCTCACCTCCAAAGCCCGGAGGAGGACGAGTTCGCGGAGCTCTAGGCT 2924
Db 2103 TCACGCTCACATCAAGGGAAC---CGATGATGATGGTGAATTTGCTGAGCTGATGTT 2159
QY 2925 TCAAGTGAACACGCTCTCATCCCGCCACCAACGCTCTCATCAATTAACGTCATCGCGG 2984
Db 2160 TCAAGTGAACAACTCTTCTGATCCCCCCCCACCACTGTGCTGTGATTAACCTGGTGTGA 2219
QY 2985 TCGTGGCGGATCTCCGACGCGCATCAACAGGATACAGTCTGGGGGCGCCCTCTTCG 3044
Db 2220 TAGTGGCTGGAGTGTGATGATATCAACAGTGGCTACCAATCATGGGTCCACTATTTCG 2279
QY 3045 GCAGCTCTTCTTGGGTCATGCTGCTCCACTCTACCCGTTCTCTCAAGGGCTCA 3104

Db 2280 GGAAGCTGTTCTTTGCAATCTGGGTGATCTCCACCTCTACCTTTCTGAGGGTCTCA 2339
QY 3105 TGGGCGCCGACAAAGACGACGCCACCGTTGTGTGTCATCTGTCATCTGCTGGCTCCA 3164
Db 2340 TGGGAAGCAGAACCGCACACCGACCATCGTCAATGTTGTCGCTCTTCTTCTTCCA 2399
QY 3165 TCTTCTCCTCTCTGCTGGGTGAGGATCGACCCCTTTCATCTGTCAGGACC---AAGGCGCGG 3221
Db 2400 TATTCTCCTCTGCTGGGTGAGATCGACCCCTTTCATATCCCTACCCAGAAGGCTCTTT 2459
QY 3222 ACCTCAGGACGATGTTGGCATCAATGCTGAGCTGTT 3256
Db 2460 CCCGTGGGACGTTGGTGTAACTGCTGAATGAT 2494

RESULT 12
CL964957
LOCUS OsIFC0011369 Oryza sativa Expressed Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.
DEFINITION CL964957
ACCESSION CL964957
VERSION CL964957.1 GI:52384602
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 2595)
AUTHORS Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M., Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L., Wong,G.K.S., Deng,X.W. and Wang,J.
TITLE An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis
JOURNAL Unpublished (2004)
COMMENT Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES
source
1. .2595
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed Library"
/note="Oryza sativa exon trapped genomic sequences"

ORIGIN
Query Match 23.0%; Score 791.2; DB 9; Length 2595;
Best Local Similarity 61.7%; Pred. No. 5.8e-146;
Matches 1546; Conservative 0; Mismatches 813; Indels 147; Gaps 12;
QY 834 AGCCGCTGTTCAGGAAGGTGTCGATCGGTGCGTCAAGAGTGAACCCGATCCGATGGTGA 893
Db 92 AGCTCTCAGCAGCAGGCTGTCCGTACCGTCCGGGAGCTCAATGATGACCGCGCGG 151
QY 894 TCGTGGTGGTCTGCTGTGTGCTCGCTTCTTCCTCCGATACCGTATCTGACCCCGGCC 953
Db 152 TCGGCTCCGCTCGTCTCATCGCGCTTCTTCCGTACCGGTACCGCGTCAACCGCGCGT 211
QY 954 CGGACGCGATCGGCTGTGCTCTGCTCCATCATCTGGAGATCTGGTTCGCCATCTCT 1013
Db 212 CCGACGCCACCGCTGTGGGTGACGTCGCTCGGTGCGAGCTCTGGCTCGCGCTCAT 271
QY 1014 GGATCTCGACAGTTCGCCAAGTGGTTCCCATCGACCGGAGACGATCTCGACCGCC 1073
Db 272 GGCTGATCGCCCACTCCGAGAGCTTCCCGCGCAACCGGCTCAGTACCTCGACAGGC 331


```

Db      2465 ACCGGGCGCGCGACCATCGCGTCTCTGGTGGTGGTCTTCTGGTCTCGGTGGGTGCG 2524
QY      3174 TGCTTGGGTGAGGATCGACCTTTTCATCGTCAGGACCAAGGCGCC 3219
Db      2525 TGCTTGGGTGAGGATCGACCTTTTCATCGTCAGGACCAAGGCGCC 2570

RESULT 13
CD726831
LOCUS   1874 bp mRNA linear EST 26-JUN-2003
DEFINITION Cucurbita pepo testa substracted cDNA Cucurbita pepo cDNA
clone CES similar to Cellulose synthase, mRNA sequence.
ACCESSION CD726831
VERSION   1
KEYWORDS EST.
SOURCE   EST.
ORGANISM Cucurbita pepo
           Cucurbita pepo
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           rosids; eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
           1 (bases 1 to 1874)
REFERENCE Bezold,T.N., Mathews,D., Loy,J.B. and Minocha,S.C.
AUTHORS Molecular analysis of the hull-less seed trait in pumpkin:
TITLE Expression profiles of cell wall related genes during development
JOURNAL Unpublished (2003)
COMMENT Contact: Subhash Minocha, Dennis Mathews, Brent Loy, Todd Bezold
           Dr. Minocha
           University of New Hampshire
           Rudman Hall, Durham, NH 03824, USA
           Tel: 603 862 3840
           Fax: 603 862 3784
           Email: sminocha@cisunix.unh.edu
           Degenerate primers and Taq were used to amplify cDNA for TOPO TA
           (Invitrogen, Carlsbad, CA) cloning. Sequencing was performed three
           times using the Dyenamic ET Terminator Sequencing kit (Amersham
           Pharmacia Biotech Inc, Piscataway, NJ). Sequences were identified
           by NCBI BLAST(X).
FEATURES             Location/Qualifiers
     source           1..1874
     -organism="Cucurbita pepo"
     -mol_type="mRNA"
     -db_xref="taxon:3663"
     -clone="CES"
     -dev_stage="20 days post-anthesis"
     /clone_lib="Cucurbita pepo testa substracted cDNA"
     /note="Organ: Testa; Total RNA was isolated from 20 day
     post-anthesis testa tissue and used in a subtraction
     hybridization procedure as according to the Clontech
     PCR-Select cDNA Subtraction kit (PT1117-1) (Clontech, Palo
     Alto, CA)."
ORIGIN
Query Match      22.7%; Score 782.4; DB 6; Length 1874;
Best Local Similarity 64.7%; Pred. No. 3.1e-144;
Matches 1213; Conservative 0; Mismatches 601; Indels 60; Gaps 1;

QY      1015 GATCTCGACCAAGTTCCTCCCAAGTGGTTCCCACTGACCGCGAGAGTACTCGACCGCT 1074
Db      1 GGTGTGGACCAAGTTCCTCCCAAGTGGTTCCCACTGCTCTGCTGAGGACACATTTATTGACAGATT 60

QY      1075 CTCCTCTAGGTACGAGAGGAGGAGCGGTGGTGTGTCGGCGGTGGACCTGTTGCT 1134
Db      61 ATCAGCCAGGTTTGAAGAGAGGAGAACCTTCCCAAGTGGTGTGCTGTGGATTTTTTGT 120

QY      1135 GAGCAGGTGGACCGCTCAAGAGCGCGGTGGTGGACCGCAACACCGTCTCCAT 1194
Db      121 GAGTACCGTGTGATCGTTGAAGAGAACCTTCTTAATCACTGCGAATACCGTGTCTCAAT 180

QY      1195 CCTCGCGGTAGACTACCGGTGGACAAGGTCTCTGCTACGTCTCCGACGAGCGGCGCTC 1254
Db      181 CTTGTGTGGACTATCTCTGATGAAGAGTCTCTGCTATGTCGATGACGCGGTGACG 240

QY      1255 GATGCTGAGTTCGAGTCGCTCTCGGAGACGCGCGAGTTTCGCGCGCAAGTGGGTGCCCTT 1314

```

```

Db      241 TATGCTCACAATTGAATCTCTAGTTGAAACAGCTACTTTGCAAGNAATGGTTTCCATT 300
QY      1315 CTGCAAGAAGTTTCGGCATCGAGCCCGCGCCCGGAGTTCTACTTCTCGCTCAAGGTGCA 1374
Db      301 CTGCAAAAAATTTCTCCATTGAACCTCGAGCTCTCTGAGTTTACTTCTCCAGNAATAGA 360
QY      1375 CTACCTCAAGGACAAAGGTGCGAGCCCACTTCGTGAGAGCGCCGCGCATGAAGAGAGA 1434
Db      361 CTATTGGAAGATAAAGTACAACTCTTTTGTAAAGGAAAGTAGAGCTATGAAGAGAGC 420
QY      1435 GTATGAGGAGTTCAGGTTCGGATCAACGCGCTGGTGGCCAGGCCATGAAGGTCCGCG 1494
Db      421 TTATGAAGAGTTTAAAGTACGAGTCAATGCTTTGGTGGCAAGGCGCAAAAAACCTCGA 480
QY      1495 AGAGGGGTGGATCATGAAGAGCGGCAACGCGCTGGCCCGGGAAACAACACCGCGACACCC 1554
Db      481 TGAAGGCTGGTCCATGCAAGATGGAACAGCTTGGCCAGGAACAATCCACGTGATCACCC 540
QY      1555 CGGCATGATCCAGGTGTTCTTGGGGCCACAGCGGGGGCCACGACACCGAGGGCAACGAGCT 1614
Db      541 TGGATGATTCAGGTTTTTCTTGGAAAATACCGCGCGCCCATGACGTAGAGGGAATGAAT 600
QY      1615 GCCCGCCTCTGTACGTCTCCGCTGAGAAGCGCGCGGATTCAGACGACCAACGAAGGC 1674
Db      601 TCCTCGACTAGTCTACGTCTCCAGAGAGAAGGCGCTGTTACCGACGATCAAAAAAGC 660
QY      1675 CGGCGCCATGAACGCTCTGATTCCGCTGCTCCGCGGTGCTACCAACGCGCCCATTCATGCT 1734
Db      661 TGGTGTGCAAAATGCTCTGGTAAGATATCTGCAGTCTCTACAATGCCCCCTTCATTTCT 720
QY      1735 CAACTTGAAGTGTGATCACTACATCAACAACAGCAAGGCCATCCGGGAGGCCATGTGCTT 1794
Db      721 TAACTTGAAGTGTGATCACTATTAACAATAGTCAAGCTATACGTAGAGGCAATGTGTT 780
QY      1795 CCTCATGACACCTCAGGTCCGCGGAGGTCTGCTACGTTCCAGTTCGCGAGAGGTTGCA 1854
Db      781 CTTGATGAGACCCACAAGTTGGCGGAGATGTATGTATGTTTCTCTCAGAGGTTTGA 840
QY      1855 CGGCATCGACGCTGCAACGACCATACGCTTAAACAGGAACACCGCTCTTCTTCGACATCAACAT 1914
Db      841 TGGCAATGATCGAAGTGTGATGATGCCAATCGCAACACGCGTTCTTCGATGTGAACAT 900
QY      1915 GAAAGGGTGTGACGCGCATCCAAAGCGCGGTGTGACGTGGGACAGGGTGTGTTCCGGCG 1974
Db      901 GAAAGGATTTGATGGCATCCAGGACACAGTTTATGTTGTTGTTGTTGTTGTTTCAATAG 960
QY      1975 CCAGGCGTCTACGCTACACCTTCCCAAGGGACCCCAAGAGGCCCAAGATGGTACCTG 2034
Db      961 ACAGGCACTTTATGGCTACGGACCTCTTCCATGCTAGCTTATCTAAGACATCTCTCATC 1020
QY      2035 CGACTGCTGCCGCTGCTTCGGCGCGCAAGAGCGGAACACACCGCAAGGACGGCTGCCGA 2094
Db      1021 ATCTTGCTCTGTTGGTGGCTGCTGCTCTTGTGCTGCTGCTCTTCCAAAGAGATCTCAAAAGA 1080
QY      2095 GGGCAACCGCTGATATGGGAGTAGATAGCGACAAGGAG-----ATGCTCATGTCACATGAATTT 2131
Db      1081 TCCGACTGAGATTGAGAGATGCAAAAAGAGAGAGAGCTTGATGCTGCAATCTTTTAACT 1140
QY      2132 -----ATGCTCATGTCACATGAATTT 2154
Db      1141 CAGGGAATAGATAATTATGATGATATGAGAGATCAATGTGATTTCTCAACTGAGCTT 1200
QY      2155 CGAGAAGCGGTTCCGGGAGTCCGGCGGTTCGTTCAGTCGACGCTGATGGAGAGGAGGGG 2214
Db      1201 TGAGAAAACCTTCGATTTGCTGTTGTTTATGAGTCTACGCTAATGGAATATGGCGG 1260
QY      2215 CGTCCCTCTCTGTCGAGCCCGCGCGCTCCTCAAGGAGGCCATCCATGTGTCAGCTG 2274
Db      1261 AGTTGCGGAATCTGCAAAATCCCTCGACTTTGATCAAGGAAGCAATTCATGTGTCATTAGCTG 1320
QY      2275 CGGCTACGAGGACAAAGACCGGACTGGGGGCTGGAGTGGGTGGATCTACGGGTGATCAC 2334

```

Db 1321 TGGTTATGAAGAGAAGTCCACTTGGGAAAGAGAGATTGGTTGGATATATGGGTCAGTGAC 1380
Qy 2335 GGAGGACATCTCTGACGGGGTTCAAGATGCACTGCGGGGGTGGCGCTCCGGTGTACTGCAT 2394
Db 1381 TGAGGATATCTTAACGGGTTTCAAGATGCACTGCGGGGGTGGAGGTCCTACTACTGCAT 1440
Qy 2395 GCCAAAGCGGCGGGTTCAGGGGTCGGGCCCGATCAATCTATCGGACGGTCTCAACCA 2454
Db 1441 GCCATTAGGCGAGCATTCAAAGGGTCGGACCAATTAACCTTTCTGATCGTCTCCACCA 1500
Qy 2455 GGTCTCGGTGGGGCTGGGGTCCGTGAGATCTTCTTACGCGGCACAGCCCCCTGCT 2514
Db 1501 AGTTCTTCGATGGGCACCTGGATCTGTGAGATTTTCTTAGCAGACACTGTCCATTATG 1560
Qy 2515 GTACGGCTTACAAGAAAGCAACCTCAAGTGGCTGGAGCGCTTCGCCCTACATCAACACAC 2574
Db 1561 GTATGATTTTCAGGGCGGCCCTCAATAGTCTCAAGAATGGCTTACATAAACCACAT 1620
Qy 2575 CATCTACCCCTTACCTCGCTCCCGCTGCTCGCTACTGCAACCTCCCGCGCTCGCT 2634
Db 1621 TGTCTATCCCTTCACTCGCTCCCTCTTGTGCTTACTGCTCAATGCTGCAATCTGCT 1680
Qy 2635 CTTACCGGCAAGTTCATCATGCGCTCGATGAGACGTTCCCGAGCCTCTTCTTCATGCG 2694
Db 1681 GCTCACAGGAAAGTTTCATTCCTCAACCGCTCTCGAAGCTTCAAGAGTACCCCTTTTCTCGG 1740
Qy 2695 CCTCTCATGTCCATCTTCGCGAGCGGCATCTCGAGATGCGGTGGAGCGGGTGAGCAT 2754
Db 1741 TCTCTTCTTGTCCATCATCTCAAGAGTCTCTGAGCTGCTTGGAGTGGTGTAGCAT 1800
Qy 2755 CGAGGAGTGTGGAGGAACGAGCAGTCTTGGGTTCATCGCGCGCGTGTCCGCGCATCTCTT 2814
Db 1801 CGAAGATATATGGGTACGAGCAATCTCGGTAATCGGAGCGCTCCGCGCACATCTCTT 1860
Qy 2815 GCGCGTGTGAGG 2828
Db 1861 TGCGGTCTTCCAAG 1874

RESULT 14
LOCUS CF430024 789 bp mRNA linear EST 03-SEP-2003
DEFINITION PH1_25_C04_g1_A002 Phosphorous-deficient seedlings Sorghum bicolor
cDNA clone PH1_25_C04_A002 5', mRNA sequence.
ACCESSION CF430024
VERSION CF430024.1 GI:34442725
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 789)
AUTHORS Cordonnier-Pratt,M.-M., Suzuki,Y., Sugano,S., Klein,R.R., Liang,C.,
Sun,F., Sullivan,R., Harris,K., Eastman,A. and Pratt,L.H.
TITLE An EST database from Sorghum: phosphorous-deficient seedlings
JOURNAL Unpublished (2003)
COMMENT Other ESTs: PH1_25_C04_b1_A002
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & M University;
sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below Phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude polyA.

Seq primer: Sug5 (CTTCTGCTCTAAAGCTGCG).
FEATURES
source 1..789
Location/Qualifiers
/organism="Sorghum bicolor"
/mol_type="mRNA"
/cultivar="BTx623"
/db_xref="taxon:4558"
/clone="PH1_25_C04_A002"
/lab_host="DH10B-TI phage-resistant E. coli"
/clone_lib="Phosphorous-deficient seedlings"
/note="Vector: pME18S-FL3; Site 1: XhoI; Site 2: XhoI; The
library was prepared from polyA+ RNA from BTx623 seedlings
grown hydroponically in the absence of added phosphorus.
At 14 days of age, roots and shoots were harvested and
stored at -80 C until RNA was isolated. Double-stranded
cDNA was cloned unidirectionally into different DraIII
sites of the pME18S-FL3 vector (5-prime DraIII site is
CATCTGTG, 3-prime DraIII site is CACCATGTG). XhoI excises
the cDNA insert."

ORIGIN

Query Match 22.0%; Score 757; DB 7; Length 789;
Best Local Similarity 97.5%; Pred. No. 2.9e-139;
Matches 769; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
Qy 2198 CTGATGGAGGAAGCGCGTCCCTCTTCTGTCGAGCCCGCGCTCTCAAGAGGCC 2257
Db 1 CTCATGGAGGAGCGCGTCCCTCTGTCGAGCCCGCGCTCTCAAGAGGCC 60
Qy 2258 ATCATGTCTATAGCTCGGCTACGAGGACAAAGCCGACTGGGGCTGGAGCTGGGTGG 2317
Db 61 ATCATGTCTATAGCTCGGCTACGAGGACAAAGCCGACTGGGACTTGGGTGG 120
Qy 2318 ATCTACGGGTGCGATCAGGAGGACATCTCGACGGGTTCGAAGATGCACTGCGCGGTGG 2377
Db 121 ATCTACGGGTGCGATCAGGAGGACATCTCGACGGGTTCGAAGATGCACTGCGCGGTGG 180
Qy 2378 CGCTCCGTGTACTGTCATGCCGAAGCGGCGCGTTCGAAGGGTTCGGCCGATCAATCTA 2437
Db 181 CGCTCCGTGTACTGTCATGCCGAAGCGGCGCGTTCGAAGGGTTCGGCCGATCAATCTA 240
Qy 2438 TCGACCGTCTCAACAGGTGCTCCGTTGGGGCTGGGGTCCGTCGAGATCTTCTTCAGC 2497
Db 241 TCCGACCGTCTCAACAGGTGCTCCGTTGGGGCTGGGGTCCGTCGAGATCTTCTTCAGC 300
Qy 2498 CGGCACAGCCCTGCTGTACGGCTACAGAAACGGCAACCTCAAGTGGCTGGAGCGCTTC 2557
Db 301 CGGCACAGCCGCTGTGTACGGCTACAGAAACGGCAACCTCAAGTGGCTGGAGCGCTTC 360
Qy 2558 GCCTTACATCAACACCATCTACCCCTTCACTCTGCTCCGCTGCTCGCTACTGCACC 2617
Db 361 GCCTTACATCAACACCATCTACCCCTTCACTCTGCTCCGCTGCTCGCTACTGTACC 420
Qy 2618 CTCGCCCGCTGCTCCCTCACCGGCAAGTTCATCATGCGTCGATTCGATTCGACCTTCGCC 2677
Db 421 CTCGCCCGCTGATGCTCCCTCACCGGCAAGTTCATCATGCGTCGATTCGATTCGACCTTCGCC 480
Qy 2678 AGCCTCTTCTTCATCGCCCTCTTTCATGTCCATCTTCGCGAGCGGCATCTCGAGATCGG 2737
Db 481 AGCTTGTCTTCATCGCGCTCTTCATGTCCATCTTCGCGAGCGGCATCTCGAGATCGG 540
Qy 2738 TGGAGCGGGGTGAGCATCGAGGAGTGTGGAGGAACGAGCAAGTTCCTGGGTTCATCGCGGC 2797
Db 541 TGGAGCGGGGTGAGCATCGAGGAGTGTGGAGGAACGAGCAAGTTCCTGGGTTCATCGCGGC 600
Qy 2798 GTGTCCGGCATCTCTTCGCGCTGTCGAGGGCTGCTCAAGTTCCTCGCGGATCCAC 2857
Db 601 GTGTCCGGCATCTCTTCGCGCTGTCGAGGGCTGCTCAAGTTCCTGGGAGGATCCAC 660
Qy 2858 ACCAATTCACCGTCACCTCCCAAGGCCAACCGCGGACGAGGACGAGTTCGCCGAGTCTC 2917
Db 661 ACCAATTCACCGTCACCTCCCAAGGCCAACCGCGGACGAGGACGAGTTCGCCGAGTCTC 720

Qy	2918	TACGCTTCAGTGGACACGCTGCTCATCCGGCGCCACCGCTGCTCATCATTAAGCTC	2977
Db	721	TACGCTTCAGTGGACACGCTCCTCATCCGGCGCCACCGCTGCTCATCATCAATC	780
Qy	2978	ATCGGCGTC	2986
Db	781	ATCGGCGTC	789

RESULT 15				
CG454544/c				
LOCUS	CG454544	943 bp	DNA linear	GSS 17-SEP-2003
DEFINITION	PUIKA96TDB_ZM_0.6_1.0_KB	Zea mays genomic clone ZMMBTa0601023,		
		genomic survey sequence.		

ACCESSION
CG454544
VERSION
CG454544.1
KEYWORDS
GSS.
SOURCE
Zea mays
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 943)	Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Benntzen, J.	Maize Genomics Consortium	Unpublished (2003)	
		Other GSSs: PUIKA96TD PUIKA96TBB		
		Contact: Cathy Whitelaw		

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

```

FEATURES
source
Location/Qualifiers
1..943
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/db_clone="ZMMBTra0601023"
/clone_lib="ZM 0.6-1.0 KB"
/notes="vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

```

ORIGIN

Query Match	21.7%	Score	747.6;	DB	9;	Length	943;
Best Local Similarity	96.9%	Pred. No.	2.2e-137				
Matches	762;	Conservative	0;	Mismatches	24;	Indels	0;
						Gaps	0;

Qy	2657	CCGTCGATTAGACAGTTGCGCCAGCCTCTTCTTCATGCGCCCTCTTTCATGTCCATCTTCGCG	2716
Db	908	CCGAGATTAGACAGTTGCGCCAGCCTCTTCTTCATGCGCCCTCTTTCATGTCCATCTTCGCG	849

Qy	2717	ACGGGCATCTCTGGAGATGCCGTGGAGCGGGGTGAGCATCGAGAGTGTGCGAGGAACGAG	2776
Db	848	ACGGGCATCTCTGGAGATGCCGTGGAGCGGGGTGAGCATCGAGAGTGTGCGAGGAACGAG	789

Qy 2777 CAGTTCTGGGTCAATCGCGCGCGGTGTCGCGCGCATCTCTTCCCGCTCGTGCAGGSCCTGCTC 2836
|||||
Db 788 CAGTTCTGGGTCAATCGCGCGCGGTGTCGCGCGCATCTCTTCCCGCTCGTGCAGGSCCTGCTC 739

QY	2837	AAGGTCTCTCCGGGGATCGACACAACTTCACCGTCACTTCAAGGCCACCGGCGACGAG	2896
Db	728	AAGGTCTCTCCGGGGATCGACACAACTTCACCGTCACTTCAAGGCCACCGGCGACGAG	669

Qy	2897	GACGACGAGTTCCCGAGCTTTCAGCCTTCAAGTGACACACGTCTCTCATCCGCCGCACC 	2956
Db	668	GACGACGAGTTCCCGAGCTTTCAGCCTTCAAGTGACACACGTCTCTCATCCGCCGCACC 	609

This Page Blank (uspto)

QY 538 rAlaValLeuThrAanAlaProPheMetLeuAnLeuAspCysAspHisTyrIleAsnAs 558
DB 1736 TGCTGTTCTCAAAATGCTCTTTTCATGTTGAACCTGGATTGTGATCACTATGTCATAATA 1795
QY 558 nSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysVa 578
DB 1796 CAGCAAGGCTGCCGAGAGGCATGCTTCTTGATGACCCACAAACTGGGAGAGGT 1855
QY 578 lCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAs 598
DB 1856 CTGCTATGTCAGATTTCTCAAAAGATTGATGATGATGATGATGATGATGATGATGATGAT 1915
QY 598 nArgAsnThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProVa 618
DB 1916 CAGGAACACAGTTTCTTTGATATTAACATGAAGGGTCTAGATGATATTAAGGTCCTGT 1975
QY 618 lTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProProLy 638
DB 1976 ATATGTGGGACTGGATGTTTTCAGAGGCAAGCTTTGATGCTATATATCTCCCAA 2035
QY 638 sGlyProLysArgProLysMetValThrCysAspCysCysProCysPheGly---ArgLy 657
DB 2036 GGGTCCAAAGCGTCCAAAAATGGTAAGCTGTGATTGTTGCCGCTGTTTGGAGCGCAA 2095
QY 657 sLysArgLysHisAlaLysAspGlyLeuProGlnGlyThrAlaAspMetGlyValAspSe 677
DB 2096 GAAGTATTAAGGAGAAGATGATGCAAAATGAGAGAGCTGCAAGCCCTAAAGGAGATGATGA 2155
QY 677 rAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAl 697
DB 2156 TGACAAAGAGGTGATGTTCCCAATGAATTTTGAAGAAGAAATTTGACCAATCTCTAT 2215
QY 697 aPheValThrSerThrLeuMetGluGluGlyGlyValProProSerSerProAlaAl 717
DB 2216 TTTTGTGACTTCTACCTTGATGGAAGAGGGTGTGCTCTCTCTTCAAGTCAGCTGC 2275
QY 717 aLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGl 737
DB 2276 CCTGCTTAAGAAGCCATTCATGTGATGATGCTGGATATGAAGATAAATCAAGTGGG 2335
QY 737 lLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMe 757
DB 2336 ACTTGAGCTTGTTGGAICTATGATGATCATCAGAGATATTCACAGGTTTAAAGT 2395
QY 757 tHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySe 777
DB 2396 GCATTGCCGTGGTGGAGTCCATTTATTGATGCAAGAGAGCTGCATTCAAGGGTAC 2455
QY 777 rAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerVa 797
DB 2456 TGCTCCTATCAACTGTGCAGATCGTCTCAACAGAGTTCCTGTTGGGCATTTGGTTCCAT 2515
QY 797 lGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLy 817
DB 2516 TGAGATTTCTTTAGTCAACATTCCTATGATGCTTATGATGCTTCAAGGAAAGAGCTAAA 2575
QY 817 sTrpLeuGluArgPheAlaTyrIleAsnThrIleTyrProPheThrSerLeuProLe 837
DB 2576 GTGGCTTGAGAGATTTGGCTATGCAACACACAACTGCTATCCATTCACCTCCCTCTCT 2635
QY 837 uLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSe 857
DB 2636 AGTTGCCTACTGTATTTCTCCAGCAGTTTGTGTTACTCACTGACAAATTCATCATGCCACC 2695
QY 857 rIleSerThrPheAlaSerLeuPheIleAlaLeuPheMetSerIlePheAlaThrGl 877
DB 2696 GATTAGCACCTTGTGCTGTTGTTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2755
QY 877 lYleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGlnPh 897
DB 2756 TATTCTTGAGTTGAATGGAGTGGAGTGAGCATTTGAGGAATGGTGGAAATGACAGTT 2815

QY 897 eTrpValIleGlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysVa 917
DB 2816 TTGGGTCAATGGTGGTGTATCAGCTCACTCTTTGCTGTTATACAGGTCTGCTAAAGGT 2875
QY 917 lLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspAs 937
DB 2876 TCTGGCTGGAATTGACACCAATTTTCACTGTTTACATCAAGGCAACA-----GATGATGA 2929
QY 937 pGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThrLe 957
DB 2930 AGAGTTTGGAGATTGTACACCTTTAAGTGGACTACACTCTTTGATTCTCCTCAACCACTAT 2989
QY 957 uLeuIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTy 977
DB 2990 TTTGATCATTAACATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3049
QY 977 rGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLe 997
DB 3050 CCAATCTGGGAGCACCTTTTGGAAAGCTCTTTTCTTCTTCTGGGTGATTGTCATCT 3109
QY 997 uTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIle 1017
DB 3110 CTATCCATTCCTTAAAGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3169
QY 1017 eTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIl 1037
DB 3170 TTGTCAGTGTATTTGCTCTATTTTCTCTTACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3229
QY 1037 eValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
DB 3230 CCTCAAGACTAAGGAGCTGTATACCAAGCTATGTGAATCAACTGC 3275
RESULT 2
AAA67114
ID AAA67114 standard; DNA; 3851 BP.
XX
AC AAA67114;
XX
DT 31-OCT-2000 (first entry)
XX
DE Pinus radiata cellulose synthase nucleotide sequence SEQ ID NO:69.
XX
KW Eucalyptus grandis; pinus radiata; Monterey pine; modification;
XX plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;
XX transgenic plant; ds.
XX
OS Pinus radiata.
XX
FN WO200022092-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-NZ000169.
XX
PR 13-OCT-1998; 98US-00170862.
PR 11-AUG-1999; 99US-0148426P.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
FI Bloksberg LN;
XX
DR WPI; 2000-339328/29.
DR P-PSDB; AAB16307.
XX
PT New genes encoding proteins involved in a plant polysaccharide
PT biosynthetic pathway, useful for modulating or altering the
PT polysaccharide content, composition or structure of the plant.
XX
PS Claim 1; Page 71-72; 301pp; English.
XX
CC The present invention describes isolated polynucleotides (PN) comprising
a sequence selected from one of 835 nucleotide sequences given in

CC ABA67073 to ABA67907, their (reverse) complements, sequences producing an
 CC Expectation (E) value of 0.01 or less compared to the 835 sequences
 CC sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the
 CC 835 sequences or sequences that are degenerately equivalent or allelic to
 CC the 835 sequences. The polynucleotides are used to modify the activity of
 CC a polypeptide involved in a polysaccharide biosynthetic pathway in the
 CC plant. They are especially used to modulate or alter the polysaccharide
 CC content, composition or structure of the plant. ABA16268 to ABA16340 are
 CC proteins encoded by some of the polynucleotide sequence given in the
 CC present invention
 XX
 SQ Sequence 3851 BP; 1040 A; 796 C; 944 G; 1071 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0 Length: 3851
 Score: 4470.00 Matches: 838
 Percent Similarity: 84.47% Conservative: 92
 Best Local Similarity: 76.11% Mismatches: 105
 Query Match: 79.14% Indels: 66
 DB: 3 Gaps: 16

US-10-627-132-30 (1-1052) x ABA67114 (1-3851)

QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgGluLeuValLeu 20
 DB 153 ATGGAACCGACGCGCGCTGGTTCGCGGTTCATACAGAACGAGTTCGTGGTCATC 212
 QY 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGluLeuCys 40
 DB 213 CATGCATGAGGAGCGGAGCGCTTGAACACGCTTGAGTGGCCACGCTGCCAGATTGT 272
 QY 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
 DB 273 GGCAGACGCTCGCGGCTTAACACAGACGCGGAGCTGTGCTGTGCTGTATGAGTGGCGG 332
 QY 61 PheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsnCysPro 80
 DB 333 TTTCCTGCTGTCGCGCGTCTATGATACAGACGAGAGAGAGAGAAATCATGCTGGTCCCG 392
 QY 81 GlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAspAsp 100
 DB 393 CAGTGCAATATCTCGTTACAGCGTCAAAAGAGGAGTCCACGGTGGAGGTGACGATGAT 452
 QY 101 GluGluAspLeuAspLeuGluHisGluPheAsnLeuAspGluAsnGlnArg 120
 DB 453 GAAGAAGACGTTGATGACATGAACATGATTTAATGTGGAGACTCAGCAAGAAACAGG 512
 QY 121 GlnLeuGluGlyAsnMetGlnAsnSerGlnLeuThrGluAlaMetLeuHisGlyArgMet 140
 DB 513 CAG-----CAGATCACCGAGGCGATGCTCCACGGAGCGCATG 548
 QY 141 SerTyrGlyArgGlyProAspAspGlyAspGlyAsn-----AsnThrProGlnIlePro 158
 DB 549 AGCTATGCGCGAGGTCCGACGACGAGAAATTCGACGATGTGCTATAATCCACAGAGCTTCT 608
 QY 159 -----ProIleLeuThrGlySerArgSerValProValSerGlyGluPheProIle 175
 DB 609 CCGCAGATTCTGTACTTGCAACCGCCACTCGGT-----GTGAGTGGGAGATTCCA--- 662
 QY 176 ThrAsnGlyTyrGlyHisGlyGluVal-----SerSerSerLeuHisLysArgIleHis 193
 DB 663 ACGTCATACTACGACACACCAATGTCTGCCAACCTCGCAATGCTGAAGCGGTGTGCAT 722
 QY 194 ProTyrProValSerGluProGlySerAlaLys-----TrpAspGluLysGlu--- 210
 DB 723 -----CCAAGTCCGAGCGGGAGTGGAGGATCATCATGATCCAAACAGAGGATATT 776
 QY 211 -----ValSerTrpLysGluArgMetAspAspTrpLysSerLys 223
 DB 777 GGTCTTATGGCTTTGGGAACGTGCTTGGAGGAGGAGCGGATGTTATTAAATCGAAG 836
 QY 224 GlnGlyIleLeuGly-----GlyGly--- 230

DB 837 GAAAAAATAATCAGGCCAGTTGGATATGACGGAAGGGAGATATCAATATATATGGGGGTTT 896
 QY 231 -----AlaAspProGluAsp---MetAspAlaAspValAlaLeuAsnAspGluAlaArg 247
 DB 897 GCACCAATATGACCGTGAAGATTATATTGATCCGATATGCCAATGACCATGACCAAGG 956
 QY 248 GlnProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetVal 267
 DB 957 CAGCCACTGTCCGAAAGTGCCTTCTTCAAGCAAAATAAATCAATCAATCGAATGGTC 1016
 QY 268 IleValValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProVal 287
 DB 1017 ATTGTAATTCGACTGATGCTGGGTATTTTCTCCGCTATGCTCTCTCATCCAGT 1076
 QY 288 ProAspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSer 307
 DB 1077 AAGATGCAATATGGGCTCTGGGCCACTTCTATGCTTTGTGAAATCTGGTTTGGCTGTGTC 1136
 QY 308 TrpIleLeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArg 327
 DB 1137 TGGATTCTTGTATCAGTTTCCCAAGTGGTTGCCCTATCAGTCGTGAAACGCTATCTGTATCGA 1196
 QY 328 LeuSerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPhe 347
 DB 1197 CTGTCAATTAAGGTATACGACGAGAGCGCAACCATCAATGCTTGACCTGTTGACCTCTTT 1256
 QY 348 ValSerThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSer 367
 DB 1257 GTGAGTACTGTATGATCCACTGAAGGAGCCCTCTTGTGTTACTGCCAATACAGTATTATCA 1316
 QY 368 IleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAla 387
 DB 1317 ATCTTTTCAGTAGACTACCTGTAGACAATGTCTCTGTTATGCTCTCTGTATGACGAGCG 1376
 QY 388 SerMetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValPro 407
 DB 1377 TCGATGCTTACTTTTGAATCTCTCTCTGAGACCTCAGAAATTTGCCAGAAATCGGTACCA 1436
 QY 408 PheCysLysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysVal 427
 DB 1437 TCTGCAAGAAATTCGACATTGAGCTTCGCGCTCCCGAAATCTATTCTCTCAGAAATTT 1496
 QY 428 AspTyrLeuLysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArg 447
 DB 1497 GACTATCTGAAGGACAAATTTCAACCCACCTTTGTCAAGAGCGCGCTGCATCAAGAGA 1556
 QY 448 GluTyrGluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValPro 467
 DB 1557 GAATATGAAGAAATTCAGGTGCGCATCAATCGGTGCTGCTGCAAGGCGCTCTAAAGTGCCC 1616
 QY 468 AlaGluGlyTrpIleMetLysAspGlyThrProTyrProGlyAsnAsnThrArgAspHis 487
 DB 1617 AAGAGAGATGGAATTCGAAGACGTTACGCTTGGCTGGTAAATAATACCCGTGACCAT 1676
 QY 488 ProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGlu 507
 DB 1677 CTTGTTATGATCCAAGTGTCTTGGGTCAAGTGGCGGCTCGATACAGAGGCAATGAG 1736
 QY 508 LeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLys 527
 DB 1737 CTTCTCCGGCTAGTATATGTTTCTCGTGAAGAGACCTGGTTTCCAGCATCAAGAG 1796
 QY 528 AlaGlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMet 547
 DB 1797 GCCGGTCCCATGAATGTTTGGTTCCGGTTTCTGTGCTGTGTCCCAATGTCTCATTTATG 1856
 QY 548 LeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCys 567
 DB 1857 CTGAATCTGATTTGTGATCACTACATTAACTAGCAAGGCAATCAGGAAGGATGTGC 1916
 QY 568 PheLeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPhe 587
 DB 1917 TTTATGATGATCCTCAGGTGGGAGAAAGTCTGTTATGTGCTCAATTTCCCTCAGAGATTC 1976

QY 588 AspGlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsn 607
 DB 1977 GATGGTATTGATCGCAATGACCGTTACGCCAATCGAAACACCGTATTCTTGTATCAAC 2036
 QY 608 MetLysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArg 627
 DB 2037 ATGAAGGCTCGATGGGAATTCAGGGCTGTATATGTGGAACTGGATGCAATGCAGA 2096
 QY 628 ArgGlnAlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThr 647
 DB 2097 AGACAAGCTCTATATGGGTATGGGCTCCCAAGGCCCAAAACGTCCTCAAGATGTGACC 2156
 QY 648 CysAspCysCysProCysPheGlyArgLysLysArgLysHisAlaLys----- 663
 DB 2157 TGTGATGTCTCCCTTGTGGGTCCTCGTAAGAAGTCTCCGAAGAAAATAAGTAGCAAG 2216
 QY 664 -----AspGlyLeuProGluGlyThrAla 671
 DB 2217 AAGAGTCAGGAAATCCAGCTCCCGCTACAAATCTGGACGGGATCAGGAAGAGTAGAA 2276
 QY 672 AspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArg 691
 DB 2277 -----GGTTATGATGACGAAGAGCAATTGTTGATGAGCACTAGACTTCGAGAAGAG 2330
 QY 692 PheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGluGlyGlyValProPro 711
 DB 2331 TTTGGCCAGTCTTACGCTTTTGTCAATCCACTCTGATGGAGAAATGGTGTTCGCA 2390
 QY 712 SerSerSerProAlaAlaLeuLysGluAlaIleHisValIleSerCysGlyTyrGlu 731
 DB 2391 ACAGCAATCCAGCTGAATGTTGAAGGAGGCTATTTCATGTCATCAGCTGTGGATATGAA 2450
 QY 732 AspLysThrAspTrpGlyLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIle 751
 DB 2451 GACAAACCGAATGGGAAAGAGCTTGGATGATCTATGGATCAGTCACAGAGCAATT 2510
 QY 752 LeuThrGlyPheLysMetHisCysArgGlyTrpArgSerValTyrCysMetProLysArg 771
 DB 2511 CTGACTGGATTCAAGATGCACACTCGAGGCTGGCGTCCATTACTGTATGCCAAGCA 2570
 QY 772 AlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArg 791
 DB 2571 CGACATTTCAAGGGTCGTCTCAATCAATCTATCAGACCGGTTTGAACAGGTTGTGGT 2630
 QY 792 TrpAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyr 811
 DB 2631 TGGGCTTTGGGATCAGTAGAATTTTCATGACAGACATTCGCCAATCTGGTATGGCTAT 2690
 QY 812 LysAsnGlyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 831
 DB 2691 ---GGGGGAGGCTCGAAATGGCTTGAAGAATTTGGCTATATCAACACCATTTGTCTATCCA 2747
 QY 832 PheThrSerLeuProLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGly 851
 DB 2748 TTCACCTCTCTTCCACTCATTGCCATTATGACACTTCCAGCGCTCAGTTTGTCTACTGCG 2807
 QY 852 LysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMet 871
 DB 2808 AAAATTGGATCCCTCAGATCAGTACTTTTGCAGTCTATTATTATAGTCTTTTTCATC 2867
 QY 872 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluTrp 891
 DB 2868 TCAATTTTGGCACTGGTATTCTGAATAGGTTGAGTGGAGTGGAGTTCGATGAAGATGG 2927
 QY 892 TrpArgAsnGluGlnPheTrpValIleGlyValSerAlaHisLeuPheAlaValVal 911
 DB 2928 TGGCGAAATGAACAGTCTCGGTTATTCGAGGGGTTCTCGACATTTTTCGAGTATT 2987
 QY 912 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 931
 DB 2988 CAAGGTCTGCTCAGGTTACTGGCAGGCAATTGATACAAATTTTCACAGTCACTGCCAAGCA 3047

QY 932 ThrGlyAspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeu 951
 DB 3048 TCA-----GATGACGGTGAGTTTGGGAACTGTATGCAATTCAAATGACACACTCTCTC 3101
 QY 952 IleProThrThrLeuLeuIleAsnValIleGlyValValAlaGlyIleSerAsp 971
 DB 3102 ATTCTCTCTACAACTCTGTTCATCAACCTTGTGGGGGTGGTGTGGCGGTAGCAGAT 3161
 QY 972 AlaIleAsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPhe 991
 DB 3162 GCAATCAACAATGATTTTCAGTCATGGGCTCTCTCTTGGGTAAAGCTTTTCTTTCGATTTC 3221
 QY 992 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1011
 DB 3222 TGGGTCAATTGTGCACCTGTATCTTCTCAAGGGTCTCATGGGCGAGCAACCGAACA 3281
 QY 1012 ProThrValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpVal 1031
 DB 3282 CCCACCATCGGTGTATTGTGCAATTCGTGGCATCTGTCTCTTTCTTTCTGGGTA 3341
 QY 1032 ArgIleAspProPheIleValValArgThrLysGlyProAspValArgGlnCysGlyIleAsn 1051
 DB 3342 AGAATTGATCTTCTTGAGTAAGTTAAAGGCCAGATACATAACAATGTGGCATCAAC 3401
 QY 1052 Cys 1052
 DB 3402 TGC 3404
 RESULT 3
 AAC83247
 ID AAC83247 standard; DNA; 7234 BP.
 AC AAC83247;
 DT 14-MAR-2001 (first entry)
 DE Cellulose synthase polynucleotide sequence SEQ ID 1.
 KW Cellulose synthase; lignin; secondary cell wall construction; wood pulp;
 XX transgenic plant; paper manufacture; ds.
 OS Arabidopsis thaliana.
 XX WO200070058-A2.
 XX 23-NOV-2000.
 XX 17-MAY-2000; 2000WO-GB001890.
 XX 18-MAY-1999; 99GB-00011379.
 XX (UTMA-) UNIV VICTORIA MANCHESTER.
 XX Turner S, Taylor N;
 XX WPI; 2001-041015/05.
 DR
 XX Cellulose synthase gene expressed during deposition of secondary cell
 PT walls in lignin-containing cells, useful for modulating expression of
 PT enzymes involved in synthesis of plant cell walls and to produce
 PT transgenic plants.
 PS
 XX Claim 3; Page 28-30; 49pp; English.
 CC This invention relates to a cellulose synthase gene expressed during the
 CC deposition of secondary cell walls in cells containing lignin. The
 CC cellulose synthase gene is useful for regulating the expression of genes
 CC specifically during secondary cell deposition in lignin containing cells.
 CC It can be used to modify the structure and cellulose content of plant
 CC secondary cell walls and to produce altered plant phenotypes specific
 CC the needs of a particular industry such as in reducing the lignin of wood
 CC pulp for paper manufacturing. A construct containing a cellulose synthase
 CC promoter sequence and a gene of interest may be used in a method for the

CC production of the product of the gene of interest in a host cell that
CC produces lignin, where the product is produced only during secondary cell
CC wall synthesis. The present sequence represents the Arabidopsis thaliana
CC cellulose synthase gene of the invention
XX
SQ Sequence 7234 BP; 2196 A; 1396 C; 1468 G; 2174 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 7234
Score: 4224.50 Matches: 856
Percent Similarity: 61.35% Conservative: 87
Best Local Similarity: 55.69% Mismatches: 90
Query Match: 74.80% Indels: 508
DB: 4 Gaps: 20

US-10-627-132-30 (1-1052) x AAC83247 (1-7234)

Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
Db 92 ATGGAGCTAGCGCGGCTTGTGCGCGGTTCTCATACCGTAATGAAGTACTGTCATT 151
Qy 21 ArgGlyHisGluAsp----- 25
Db 152 CACAACCATGAAGAGGTTTTTTCACATTACTTTTTTCTCATCCTACCTACCAAGTTTTTTTT 211
Qy 25 ----- 25
Db 212 TACCAACGCTAGTAAATAATATTATGCAATTTTTTCGTTTTATTTGGTACTATTTCTTAAA 271
Qy 25 ----- 25
Db 272 AATTCGTATGTTTTGGGATAAATAATATGATCATTTTTTAAATCACTTCTTATTATGA 331
Qy 25 ----- 25
Db 332 GACAAAATTTATAATCTGTATTCTGTAGTTGCAATAATGTTGTAGAAAATTCATATCTT 391
Qy 25 ----- 25
Db 392 TGTAGCAACATATAATATTTTGGTAAATTAATTAAGTTGAGAGTCAGGTTTAACCAATT 451
Qy 25 ----- 25
Db 452 TTAATCGCTGTCATTTTTTTTATTATCTTTACTTCAATTAGAATCTTTTTTTTGTGATTAG 511
Qy 26 -----Pro 26
Db 512 AAATTTACAGGTTTTTATTTTCGTCCTTTAAGGAACCTTAAACTTTTGTGTAATTATAACAGCCA 571
Qy 27 LysProLeuArgAlaLeuSerGlyGlnValCysGluLeuCysGlyAspGluValGlyLeu 46
Db 572 AAGCCTCTGAAGAATCTAGATGGACAAATCTGTGAGATATGTGGAGATCAGATTGGTTTA 631
Qy 47 ThrValAspGlyAspLeuPheValAlaCysAsnGluCysGlyPheProValCysArgPro 66
Db 632 ACAGTAGAAGGAGACCTCTTCGTAGCTTGCATGAGTGTGTTTTCCGCGGTGTAGACT 691
Qy 67 CysTyrGluTyrGluArgGluGlyThrGlnAsnCysProGlnCysGlyThrArgTyr 86
Db 692 TGCTATGATACGAGAGAAGAGAAACACAAAACCTGCTCCTCAGTGAAGACTCGTTAC 751
Qy 87 LysArgLeuLys----- 90
Db 752 AAGCGTCTCAGAGGTAAGTATTATTATTAATCTCCCTCTGCTCTGTGTTGTCGACGAAA 811
Qy 90 ----- 90
Db 812 TGCCCTCATGAAATTTAAABAGGCTGTTCTTTTTTTTAGTTTGAATTCGAGAGTAATG 871
Qy 91 -----GlySerProArgValAlaGlyAspAspGluGluAsp 103
Db 872 ATCTGTTTTTGGTTTCTGAAGGAAGCCCAAGAGTGGAGGAGATGAAGACCAAGAAGAT 931

Qy 104 IleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnArgGlnLeuGlu 123
Db 932 ATTGATGATATTGAGTATGAATTTAATATCGAACATGAACAAGAT----- 976
Qy 124 GlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMetSerTyrGly 143
Db 977 -----AAGCATAAAGCATTTCTGCTGAGGCTATGCTTTATGGAATAATGACTATGGA 1027
Qy 144 ArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIleProIleIleThrGly 163
Db 1028 AGAGGCTCTGAGATGATGAGAAATGGG-----AGATCCCACTGTTATAGCTGGT 1078
Qy 164 SerArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrGlyHisGlyGlu 183
Db 1079 GGTCA-----AGTGGAGAAATTTCCAGTTGGAGAGGTTTATGTTAATGGAGAA 1126
Qy 184 ValSerSerSerLeuHisLysArgIleHisProTyrProValSerGluProGlySerAla 203
Db 1127 -----CATGGGCTTCATAGCGTGTGCACCCATATCCATCTCTGAAGCTGGTGAAGTCT 1180
Qy 204 LysTyrAsp-----GluLysLysGluValSer----- 212
Db 1181 CA-TGGAAATGTTAACTTACATATAGATTTAAGAAATGCTCAGAGTATGATTAGTAGG 1239
Qy 212 ----- 212
Db 1240 GTCATGCATATCTCCATATGTGCAAATAACATAAGTATGAGGCTTCCAGCTTAATAGTA 1299
Qy 212 ----- 212
Db 1300 GATAGGACATAGTTTTCATAAACATGAGACTTTGGGTTCTATTACATCTTCTTCTATGAAAT 1359
Qy 212 ----- 212
Db 1360 TCATCAGACAGACCTTTTCTAAACATACCTTCTCTTTGTTTATGTTGTGTAATTTAAT 1419
Qy 213 -----TrpLysGluArgMetAspAspTrpLysSerLysGlnGly 225
Db 1420 GTGGTAGGAGTAGGAGGAGGATGGCGGAAAGAAATGATGACTGGAAGCTCCAGCATGGA 1479
Qy 226 IleLeuGlyGlyGlyAlaAspProGlu---AspMet-AspAlaAspValAlaLeuAsn-- 243
Db 1480 AATCTTGGCGCAGAACACAGATGATGATCTCGATGGGACTGTAATGCCCTCCACAAACAT 1539
Qy 243 ----- 243
Db 1540 TTATCTAAGACATCAGTTTTTGTATGATTGATTCGCTTACAAAATTTTGGATTTCAC 1599
Qy 244 -----AspGluAlaArgGlnProLeuSerArgLysValSerIleAl 257
Db 1600 TGGAAATGGCTGTAGGATCGACGAGGCGACGCGCAGCTCTCGCGGAAAGTTCCCAITGC 1659
Qy 257 aSerSerLysValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPh 277
Db 1660 CTCAAGCAAGATCAATCCATATCGATCGGTGTCATCGTTGCTAGGCTTGTGATTCTAGCAGT 1719
Qy 277 ePheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSe 297
Db 1720 TTTTCTCGGTTATAGGCTCTTGAATCAGATGCATGCTCTGGGATATAGGCTGACCTC 1779
Qy 297 rIleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPh 317
Db 1780 TGTGATCTGTAATCTGGTTTCGCTGCTCTTGGATTCTTGGATCAGTATCCCAAGTGGTT 1839
Qy 317 eProIleAspArgGluThrTyrLeuAspArgLeuSerIle----- 330
Db 1840 CCCTATTGAACGTGAGACCTATCTAGATCGGCTTTCTCCCTCAGGTAAAAATCCACAGATTCT 1899
Qy 330 ----- 330
Db 1900 CAAGTAGAAGTCTTAAATCTATCAGCTGGAGTTTGGATGTAATATTTTGTATTATA 1959
Qy 331 -ArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSerTh 350


```
Qy 897 heTTPValIleGlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysv 917
Db 4177 TCTGGGTCATTGGAGGAATCTCAGCTCATCTCTTGGGTGTCAGAGGCTCTCTCAAA 4236
Qy 917 alLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspA 937
Db 4237 TCTTAGCAGGCAATGACACAACTTCCCGTTCACATCAAGGCAACA-----GATGATG 4290
Qy 937 spGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThrL 957
Db 4291 ATGACTTTGGAGAACTTTACGCATTCAAAATGGACACACTGCTGATCCCTCCCAACAAC 4350
Qy 957 euLeuIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnGlyLT 977
Db 4351 TCTTAATCATAAACATTTGGCGTGTGGTGGAGTCTCAGATGCAATTAACATGAT 4410
Qy 977 YrGlnSerTrpGlyProLeuPheGlyLysLeuPheAlaPheTrpValIleValHisL 997
Db 4411 ATCAGTCTTGGGACCTCTATTGTGTAACCTTCTCTCTCTCTTGGGTCAATGTTTCATC 4470
Qy 997 euTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValI 1017
Db 4471 TCTACCCATTTCTCAAGGCTCTGATGGGTAGACAGAACAGAACCAACCACTTGTGTGA 4530
Qy 1017 leTTPSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheI 1037
Db 4531 TTTGGTCAAGTGTATTGGCATCTATCTCTCTCTTGTCTTGGGTAAAGATTAATGCTTTTG 4590
Qy 1037 leValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 4591 TGCTCAAGACCAAGGACCTGACACTTCCAAGTGTGGCATCACTGC 4637

RESULT 4
AAZ58263
ID AAZ58263 standard; cDNA; 3776 BP.
XX
AC AAZ58263;
DT 08-MAY-2000 (first entry)
XX
DE Corn cellulose synthase cDNA.
XX
KW Corn; maize; cellulose synthase; transgenic plant; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT 2..3448
FT /*tag= a
FT /partial
XX
XX WO200004166-A2.
XX
XX 27-JAN-2000.
XX
XX 13-JUL-1999; 99WO-US015871.
XX
XX 14-JUL-1998; 98US-0092844P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX Allen SM, Fader GW, Falco SC, Kinney AJ, Lightner JE, Miao G;
PI Rafalski JA, Thorpe CJ;
XX
XX WPI; 2000-182431/16.
XX
XX P-PSDB; AAY58832.
XX
XX Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as
XX probes for isolating cDNAs and genes encoding homologous proteins, for
XX producing transgenic plants.
XX
XX Claim 9; Page 37-38; 93pp; English.
```

```
XX
CC The present sequence is that of a contig of cDNA clones and encodes a
CC portion (see AAY58832) of corn cellulose synthase (CS). The cDNA clones
CC were isolated from corn cob, developing tassel, seedling, shoot culture
CC and leaf tissue cDNA libraries on the basis of homology to Arabidopsis
CC and cotton CS sequences. The invention relates to isolated nucleic acid
CC fragments encoding plant CS and to CS polypeptides. It also relates to
CC the construction of a chimeric gene encoding all or a portion of the CS,
CC in sense or antisense orientation, where expression of the gene results
CC in altered levels of the CS in transformed host cells. The host cells can
CC be used to screen compounds for their ability to inhibit CS activity. CS
CC nucleic acids are also useful for producing transgenic plants having
CC altered levels of CS, and hence altered levels of fibre. CS may also
CC serve as a target for the development of novel herbicides
XX
SQ Sequence 3776 BP; 931 A; 906 C; 999 G; 940 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3776
Score: 4091.00 Matches: 760
Percent Similarity: 79.64% Conservative: 128
Best Local Similarity: 68.16% Mismatches: 137
Query Match: 72.43% Indels: 90
DB: 3 Gaps: 16

US-10-627-132-30 (1-1052) x AAZ58263 (1-3776)
Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
Db 179 ATGGAGCCAGCGCGGCTGGTGGCGGCTCGCACAAACCGAACGAGCTCGTGTATC 238
Qy 21 Arg-----GlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGlu 38
Db 239 CGCGCGATGGCGAGCCAGCGCGGCGAAGCCATGGACCGAGCGGAGCGAGGTGTGCAG 298
Qy 39 IleCysGlyAspGluValGlyLeuThrValAspGlyLeuPheValAlaCysAsnGlu 58
Db 299 ATTTGCGCGACGACGTGGGGCGCAACCCGACGGGAGCGGCTTGTGTGGGTGTCAACGAG 358
Qy 59 CysGlyPheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsn 78
Db 359 TGGCGCTTCCCATCTCTCCGGGACTGTCTAGAGTACGAGCGCGCGGAGGCGGAGGAG 418
Qy 79 CysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAsp 98
Db 419 TGCCCCCAGTGCAAGACCCCGCTTCAAGCGCTCAAGGGGTGGCGCGCGTCCCGGGGAC 478
Qy 99 AspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGln 118
Db 479 GAGGAGGAGGACGGCGTGCACGACCTGGAGAACCGAGTTCAACTGGAGCGCACACGAC 538
Qy 119 GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
Db 539 TCCAGTAGTACCTC-----GCCGAGTCCATGCTCAACGCC 571
Qy 139 ArgMetSerTyrGlyArgGlyProAsp---AspGlyAspGlyAsnAsnThrProGlnIle 157
Db 572 CACATGAGCTACGCCCGCGCGCGCGCTCGACGCGCTGCGCGAGCCATTCCACCCCATC 631
Qy 158 ProProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsn 177
Db 632 CCC-----AATGTTCCCTC-----CTCACCAC 655
Qy 178 GlyTyrGlyHisGlyGluValSerSerSerLeuHis----- 189
Db 656 GGACAGATGTCGATGATCATCCCCCGGACGACGCGCCCTGTGCTCGTGGGT 715
Qy 190 -----LysArgIleHisProTyrProValSerGluProGly-----Ser 202
Db 716 GGCGGGGGAAGAGAGATTACCTCTCCGTACGCGGATCCCAACCTTCCTGTGCAACCG 775
Qy 203 AlaLysTrpAspGluLysLysGlu-----ValSerTrpLys 214
```

Db 776 AGGCTATGACCCCTCCAAAGGATCTCGCCGATATGGCTACGGAGCGTACCATGGAG 835
Qy 215 GluArgMetAspAspTrpLysSerLysGlnGlyIleLeu-----Gly 228
Db 836 GAGAGATGAGAGCTGGAAGCAGAAAGCAGAGGAGAGTACCACGACGAGGAACGATGCG 895
Qy 229 GlyGlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGln 248
Db 896 GCGCGCGATGATGGTGAT-----GATGCAGATCTACCACTAATGATGATGAAGCTAGACAG 949
Qy 249 ProLeuSerArgLysValSerIleAlaSerLysValAsnProTyrArgMetValIle 268
Db 950 CCAATTGTCAGAAAGATCCCGCTTCTTCAAGCCAAATCAACCCCTATAGATGATATA 1009
Qy 269 ValValArgLeuValValAlaPhePheLeuArgTyrArgIleLeuHisProValPro 288
Db 1010 ATAATTGCGCTAGTGGTTGTGTTCTTCTTCCACTACCGAGTGCATCGCGTGCT 1069
Qy 289 AspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrp 308
Db 1070 GATGCATTTGCTTTATGGCTCATATCTGTGATCTGTGAATTTGTTTGGCCATGTCTTGG 1129
Qy 309 IleLeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeu 328
Db 1130 ATTCTTGACCAAGTTCCTTCTATCGAGAGGAACCTATCTTGACCGGCTG 1189
Qy 329 SerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheVal 348
Db 1190 AGTTTAAAGTTTGACAAGAGGCGATCTTCTCAACTCGCCCTGTTGATTTCTTTGTC 1249
Qy 349 SerThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIle 368
Db 1250 AGTACGGTGTGATCCCTTGAAGGAACCTCCATTTGGTCACTGCTAACTACTTCTATCTATC 1309
Qy 369 LeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSer 388
Db 1310 CTTTCGGTGGATATCCAGTTGATTAAGTTTCATGCTACGTTTCATGATGGTCTGCC 1369
Qy 389 MetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPhe 408
Db 1370 ATGCTGACATTTGAAGCATTTGCTGAACATCTGAATTTGCAAGAAATGGTTCCTTTC 1429
Qy 409 CysLysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAsp 428
Db 1430 TGCAAAAGATATAGCCTTGGAGCTCGTCTCCAGAGTGGTACTTCCACAGAAAGATAGAC 1489
Qy 429 TyrLeuLysAspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGlu 448
Db 1490 TACCTGAAGACACAGGTGGCGCAAACTTTGTTAGAGAACGGAGGCAATGAAGAGAG 1549
Qy 449 TyrGluGluPheLysValArgIleAlaLeuValAlaLysAlaMetLysValProAla 468
Db 1550 TATGAGGAATTCAGGTCAGATCAATGCTTTGGTTGCTTAAAGCCCAAAAGGTTCTCTGAG 1609
Qy 469 GluGlyTrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisPro 488
Db 1610 GAAGGATGCAATCAGCATGGAATCTCATGSCCGGAAATAATGTCGCGTATCATCCT 1669
Qy 489 GlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyValGlnLeu 508
Db 1670 GGAATGATTCAGGTTTCTTGGTCAAGTGTGGCCATGATGTGGAAGGAATGAGCTG 1729
Qy 509 ProArgLeuValTyrValSerArgLysArgProGlyPheGlnHisIleLysLysAla 528
Db 1730 CCTCGATTTGTTTATGTTTCAAGAGAAAAAGCGCCAGGCTACAAACCATCAAGAAGGCT 1789
Qy 529 GlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeu 548
Db 1790 GGTGTATGAATGATGCTGCGAGTCTCTGCTGTAATACTAACTGCTCTTATTTGCTG 1849
Qy 549 AsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPhe 568
Db 1850 AACTTGGATTTGATCATATATCAATTAATAGTAGGCTATTAAGGAAGCAATGTGTTT 1909

Qy 569 LeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAsp 588
Db 1910 ATGATGGATCTTGTCTTGGAAAGAAAGTTGCTATGTATGTCAGTTTCTCTCAAAGATTGAT 1969
Qy 589 GlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMet 608
Db 1970 GGGATTGATGCCCATGATCGATATGCTAACAGAAATGTTGTCTTTTTCGATATCAACATG 2029
Qy 609 LysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArg 628
Db 2030 AAAGTTTGGATGGTATCCAGGGCCCATTTATTTGGGTACTTGGATGTGTCTTCAGAAG 2089
Qy 629 GlnAlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCys 648
Db 2090 CAGGCATTTATATGGCTAGCTGCCAAA---ACAAAGAAAGCCACCATCAAGAACTTGC 2146
Qy 649 Asp-----CysCysProCysPheGly---ArgLysLysArgLys 660
Db 2147 AACTGCTGGCCAAAGTGGTGCATTTGCTGTGTTTTTGGTAAACAGGAAGCAAGAAG 2206
Qy 661 HisAlaLysAspGlyLeuProGlu----- 668
Db 2207 AAGACCAAGACCTCTAAACCTTAATTTGAGAAGATAAAGAACTTTTAAAGAAAAGNA 2266
Qy 669 -----GlyThrAlaAspMetGlyValAspSer 677
Db 2267 AATCAAGCCCTGCATATGCTCTTGGTGAAATTTGATGAAGCGCTCCAGGAGCTGAAAT 2326
Qy 678 AspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaIle 697
Db 2327 GAAAGGCGTATGTTGTAATCAACAGAAAGTTGGAAAAGAAATTTGGCCAGCTTTCAGTT 2386
Qy 698 PheValThrSerThrLeuMetGluGlyGlyValProProSerSerSerProAlaIle 717
Db 2387 TTTGTTGCATCCACACTTCTTGAGAAATGTTGGAAAGTCCGAGAGTCCAGCTTCT 2446
Qy 718 LeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGly 737
Db 2447 CTTCTGAAGGAAGCTATACATGTCATGTTGATGAAGACAAACAGGCTGGGA 2506
Qy 738 LeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMet 757
Db 2507 AAAGATATTTGTTGGATTTATGGATCATGTCACAGAAGATATTTCTTACCTGGGTTAAGATG 2566
Qy 758 HisCysArgGlyTyrArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySer 777
Db 2567 CACTGCCATGTTGGCGTCAATTTACTGCATACCTAAACCGGCGCCTTCAAAGTTCC 2626
Qy 778 AlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerVal 797
Db 2627 GCACCTCTCAATCTTTCGATCGTCTTCCAGGTTCTTCGTTGGGCTCTTGGTTCAAT 2686
Qy 798 GluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLys 817
Db 2687 GAAATTTTCTTCAGCAACCACTGCCCTCTCTGGTATGGGTAT---GGTGGTGCAGTAAG 2743
Qy 818 TrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeu 837
Db 2744 TTCCTGGAAGTTTTCGTACATTAATCCATCGTATACCTTGGACATCTATCCCGCTC 2803
Qy 838 LeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSer 857
Db 2804 TTGGCTATTGCATATGCTGCCATCTGCTTGTGTCAGCAGGAAATTTATCACGCCAGAG 2863
Qy 858 IleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGly 877
Db 2864 CTTAAACATGTTGGCAGCCTCTGTTGTCATGTCACCTTTTCTGCTGCTGCTGCTGCTG 2923
Qy 878 IleLeuGluMetArgTrpSerGlyValSerIleGluGluTyrTrpArgAsnGluGlnPhe 897
Db 2924 ATCTGGAATGAGATGGGTGGTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2983

```
Qy 898 TrpValIleGlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuVal 917
Db 2984 TGGGTCAATTGGAGCGGTCTTCAACATCTTTGCTGTGTCAGGACCTCTCAAGGTC 3043
Qy 918 LeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspAsp 937
Db 3044 ATAGCTGGTGTAGACACGAGCTTCACTGTGACATCCAAAG-----GGCGGAGACGAG 3097
Qy 938 GluPheAlaGluLeuValAlaPheLysTrpThrLeuLeuLeuLeuProProThrThrLeu 957
Db 3098 GAGTCTCTCAGAGCTGTACACATTCAAATGGACGACCTTCTGATACCTCCGACAAACCTG 3157
Qy 958 LeuIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyr 977
Db 3158 CTCTACTGAACTTCAATGGAGTGTAGCTGGCATCTCCAAATGCGATCAACACGGATAT 3217
Qy 978 GlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeu 997
Db 3218 GAATCATGGGGCCCTGTTCCGGGAAGCTCTTCTTTGCAATTTGGGTGATCGTCCATCTT 3277
Qy 998 TyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIle 1017
Db 3278 TACCGTTCCTCAAGGGTCTGGTTGGAGGCGAGAACAGGACGCCAACGATTTGTCATTGTC 3337
Qy 1018 TrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIle 1037
Db 3338 TGGTCCATCTCTTGGCTTCGATCTTTCGGTCTGTTTGGGTCCGATCGACCCCTTCTT 3397
Qy 1038 ValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3398 GCGAAGGATGATGTCCTCTGGAGGAGTGTTGGTCTGGATTGC 3442
RESULT 5
ID AAZ58266
AC AAZ58266 standard; cDNA; 3936 BP.
AC AAZ58266;
XX 08-MAY-2000 (first entry)
XX Corn cellulose synthase cDNA clone p0122.chkamh70rc.
XX Corn; maize; cellulose synthase; transgenic plant; ss.
XX Zea mays.
XX Key Location/Qualifiers
XX 144..3404
XX CDS /*tag= a
XX
XX WO200004166-A2.
XX
XX 27-JAN-2000.
XX
XX 13-JUL-1999; 99WO-US015871.
XX
XX 14-JUL-1998; 98US-0092844P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G;
XX Rafalski JA, Thorpe CJ;
XX
XX WPI; 2000-182431/16.
XX P-PSDB; AAY58835.
XX
XX Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as
XX probes for isolating cDNAs and genes encoding homologous proteins, for
XX producing transgenic plants.
XX
XX Claim 9; Page 49-50; 93pp; English.
XX
XX The present sequence is that of full-length cDNA clone p0122.chkamh70rc
CC
```

```
CC encoding corn cellulose synthase (CS, see AAY58835). The cDNA clone was
CC isolated from a cDNA library produced from corn pith tissue collected
CC from internode subtending ear node 5 days after pollination. The
CC invention relates to isolated nucleic acid fragments encoding plant CS
CC and to CS polypeptides. It also relates to the construction of a chimeric
CC gene encoding all or a portion of the CS, in sense or antisense
CC orientation, where expression of the gene results in altered levels of
CC the CS in transformed host cells. The host cells can be used to screen
CC compounds for their ability to inhibit CS activity. CS nucleic acids are
CC also useful for producing transgenic plants having altered levels of CS,
CC and hence altered levels of fibre. CS may also serve as a target for the
CC development of novel herbicides
CC
SQ Sequence 3936 BP; 989 A; 838 C; 1018 G; 1091 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3936
Score: 4073.50 Matches: 757
Percent Similarity: 79.68% Conservative: 129
Best Local Similarity: 68.08% Mismatches: 139
Query Match: 72.12% Indels: 87
DB: 3 Gaps: 15

US-10-627-132-30 (1-1052) x AAZ58266 (1-3936)
Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
Db 144 ATGAGCGAGAGCCCGGGCTGGTGGCGGCTCCACACACGACGAGCTCGTCTGTCATC 203
Qy 21 Arg-----GlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGlu 38
Db 204 CGCGCGGACGCGCATCCCGGGCCGAAAGCCGCCGGAGCAGAACGGGCGAGTGTGCCAG 263
Qy 39 IleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGlu 58
Db 264 ATTTGGCGGACGACGCTCGGCCCTTGCCCCCGGGGGGACCCCTTCGTGGGTGGTCAACAG 323
Qy 59 CysGlyPheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsn 78
Db 324 TGGCCTTCCCGCTCTGCCGGGACTGCTAGATACGAGCGCGGGAGGCGCACGACAGAC 383
Qy 79 CysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAsp 98
Db 384 TGCCCCCAGTGCAAGACTCGATACAGCGCTCAAGGGCTGCCAACCGTGTACCGGTGAC 443
Qy 99 AspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGln 118
Db 444 GAGGAGGAGGACGGCGTGGATGACCTGGACACACGAGTTCACCTGGGACGGCCATGACTCG 503
Qy 119 GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
Db 504 CAG-----TCTGTGGCGGAGTCCATGCTCTACGGC 533
Qy 139 ArgMetSerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIlePro 158
Db 534 CACATGAGCTACGCGCGTGA-----GGTGACCTAATGGCGGCCCAACAGCTTTC 584
Qy 159 ProIleIleThrGlySerArgSerValProValSerGlyGluPhePheIleThrAsnGly 178
Db 585 CAGCTC-----AAACCCCAATGTTCCACTC-----CTCACCACACGGG 620
Qy 179 TyrGlyHisGlyGluValSerSerSerLeuHis----- 189
Db 621 CAATGTGTGATGATCCACCGGACGACGCGGTGGTGGCTTCTTTCATGGTGGT 680
Qy 190 -----LysArgIleHisProTyrProValSerGluProGly-----SerIle 203
Db 681 GGGGAAAGAGCATACATCCCTTCTTATGCGGATCCACGCTTACCTGTGCAACCCAGG 740
Qy 204 LysTrpAspGluLysLysGlu-----ValSerTrpLysGlu 215
Db 741 TCTATGGACCCATCAAGGATCTTGTGATATGGGTATGGTATGTTGTTGGTAAGAA 800
```

QY 216 ArgMetAspAspTrpLysSerLysGln-----GlyLeuLeuGlyGly 229
DB 801 CGGATGGAGAAATGGGAAGCAGACAGAGAGAGAGATGCACACGCGGGAATGATGGTGT 860
QY 230 GlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnPro 249
DB 861 GGTGATGGTGGAC-----GATGCTGATCTACCACTAATGGATGAAGCAGACACACNA 914
QY 250 LeuSerArgLysValSerLleAlaSerSerLysValAsnProTyArgMetValLleVal 269
DB 915 CTGTCAGGAAAATCCACTTCATCAAGCCAGAGATTAATCCATATAGGATGATTAATCAT 974
QY 270 ValArgLeuValValAlaAlaPhePheLeuArgTyArgLleLeuHisProValProAsp 289
DB 975 ATTCGGCTTGTGGTGTGGGTCTCTTCCACTACCGAGTGCATCCGGTGAATGAT 1034
QY 290 AlaIleGlyLeuTrpLeuValSerLleLeuCysGluLeuTrpPheAlaIleSerTrpIle 309
DB 1035 GCATTTGCTTGTGGCTCATATCTGTATCTGTGAAATCTGGTTGCCATGTCCTGGATT 1094
QY 310 LeuAspGlnPheProLysTrpPheProLleAspArgGluThrTyLeuAspArgLeuSer 329
DB 1095 CTTGATCAATTCCTCAAGTGGTTCCTATTAGAGAGAGACTTACTAGACCGGTGTCA 1154
QY 330 LeuArgTyArgGluArgGluGlyClnProSerLeuLeuSerAlaValAspLeuPheValSer 349
DB 1155 CTGAGGTTGCGACAGAGAGGCGAGCATCTCAACTTGGCTCCAAATTCATTTCTTCAGT 1214
QY 350 ThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerLleLeu 369
DB 1215 ACGGTTGATCCCTTAAGGAACCTCTTGGTGCACAACTACTGTTCTATCTATCTTCCT 1274
QY 370 AlaValAspTyProValAspLysValSerCysTyTrValSerAspAspGlyAlaSerMet 389
DB 1275 TCGGTGGATATCTCTGTTGATAAGGTCTCTGCTATGTTCTCGATGGTGTCTCAATG 1334
QY 390 LeuThrPheGluSerLeuSerClnThrAlaGluPheAlaArgLysTrpValProPheCys 409
DB 1335 CTACGTTTGAAGCAATATCTGAACATCTGAAATTTGCAAGAAATGGGTCTCTTCTGC 1394
QY 410 LysLysPheGlyIleGluProArgAlaProGluPheTyPheSerLeuLysValAspTyTr 429
DB 1395 AAACGGTACATATTTGAACCTCGCTCCAGAGTGGTACTTCCACAGAGATAGACTAC 1454
QY 430 LeuLysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArgGluTyTr 449
DB 1455 TTGAAGACAAAGTGGCAGCAAACTTTGTTAGGAGAGAGAGCAATGAAGAGAGATAT 1514
QY 450 GluGluPheLysValArgLleAsnAlaLeuValAlaLysAlaMetLysValProAlaGlu 469
DB 1515 GAGGAATTTCAAGGTGAGAAATCAATGCCTTAGTTGCCAAAGGCCAGAAAGTTCTCTGAAGAA 1574
QY 470 GlyTrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly 489
DB 1575 GGATGGACATGCAAGATGGAAACCCCTGGCTCGGAAACAATGTTCTGGATCATCTCTGGA 1634
QY 490 MetLleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuPro 509
DB 1635 ATGATTCAGTCTCTCTTGGCCAAAGCGAGGSCCTTGACTGTGAGGGAATGAATGACGCA 1694
QY 510 ArgLeuValTyValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGly 529
DB 1695 CGATTTGGTTTATGTTTCTAGAGAGAAACGACCGAGCTATAACATCATTAAGAAGCTGGT 1754
QY 530 AlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsn 549
DB 1755 GCTATGANTGANTGGTCCGAGTCTCTGCTGTACTTAACAATGCTCCATATTTGTTAATC 1814
QY 550 LeuAspCysAspHisTyTrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeu 569
DB 1815 TTGGATTGTGATCACTACATCAACAACAGAGGCTATAAAGGAAGCAATGTGTTTATG 1874
QY 570 MetAspProGlnValGlyArgLysValCysTyTrValGlnPheProGlnArgPheAspGly 589

DB 1875 ATGGACCCCTTACTAGGAAGAAGGTTTGCTATGTACAGTTCCCTCAAGATTTGATGGG 1934
QY 590 IleAspValHisAspArgTyTrAlaAsnArgAsnThrValPhePheAspTrleAsnMetLys 609
DB 1935 ATTGATCCCATGACCGATATGCTAACCGGAATGTGTCTTTTTTGATATCAACATGAAA 1994
QY 610 GlyLeuAspGlyIleGlnGlyProValTyValGlyThrGlyCysValPheArgArgGln 629
DB 1995 GGTTTGGATGTTATTCAGGGTCCAATTTATGTTGTGCTACTGGATGTGTATTAGAGGCGAG 2054
QY 630 AlaIleuTyGlyTyArgProLysGlyProLysArgProLysMetValThrCysAsp 649
DB 2055 GCATTTATATGTTATGATGCTCCCAAA-----ACAAAGAGACCCACCATCAAGGACTTGCAC 2111
QY 650 -----CysCysProCysPheGlyArgLysLysArgLysHisAla 662
DB 2112 TGCTGGCCCAAGTGTGTGCTTTGCTGTGCTTGTGCAATAGGAAGCAAAAGAAGACT 2171
QY 663 LysAspGlyLeuProGlu----- 668
DB 2172 ACCAAACCAAAACAGAGAGAAAGTTATTTTCAAGAAAGAGAGAACCAATCC 2231
QY 669 -----GlyThrAlaAspMetGlyValAspSerAspLysGlu 680
DB 2232 CTGCATATGCTCTTGGTGAATTTGACGAAGCTGCTCCAGGAGCTGAGAAATGAAAGGCC 2291
QY 681 MetLeuMetSerHisMetAsnPheClnLysArgPheGlyGlnSerAlaAlaPheValThr 700
DB 2292 GGTATGTAAATCAACAAAATTTAGAAAGAAATTTGGCCAATCTTCTGTTTTGTTACA 2351
QY 701 SerThrLeuMetGluGluGlyValProProSerSerProAlaAlaLeuLeuLys 720
DB 2352 TCCACACTTCTCGAAGATGGTGAACCTTGAAGAGTGCAGCTCTGCTTCTCTTTTGA 2411
QY 721 GluAlaIleHisValIleSerCysGlyTyArgLysLysThrAspTrpGlyLeuGluLeu 740
DB 2412 GAAGCTATACATGTCATTTAGTTGTGGTTATGAAGACAGACAGACTGGGGGAAAAGAGAT 2471
QY 741 GlyTrpIleTyGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg 760
DB 2472 GGCTGGATCATGATGATCATGTTACAGAGATATTTCTACTGGTTTCAAGATGATGATGTCAT 2531
QY 761 GlyTrpArgSerValTyCysMetProLysArgAlaAlaPheLysGlySerAlaProIle 780
DB 2532 GGTGGGCGTCAATTTACTGCATACCTAAACGGGTGTCATTCAAAGTTCTGCACCTCTG 2591
QY 781 AsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePhe 800
DB 2592 AATCTTTCAGATCGTCTTCCACAGGTGCTTCGGTGGGCTCTTGGGTCTTATTGAGATCTTC 2651
QY 801 PheSerArgHisSerProLeuLeuTyArgTyTrLysAsnGlyAsnLeuLysTrpLeuGlu 820
DB 2652 TTCAGCATCATGCTCCCTCTTTGGTATGGGTAT---GGTGGCGGTCTCGAAATTTTGGAA 2708
QY 821 ArgPheAlaTyTrIleAsnThrIleTyTrProPheThrSerLeuProLeuLeuAlaTyTr 840
DB 2709 AGATTTCTCATCACTCACTCCATCGTGTATCTTGGACATCTATTCCCTCTTGGCTTAC 2768
QY 841 CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr 860
DB 2769 TGTCATTTGCTGCATCTGTTTATTGACGGGAAATTTATCACTCCAGAGCTGAATAAT 2828
QY 861 PheAlaSerLeuPheIleAlaLeuPheMetSerIlePheAlaThrGlyLleLeuGlu 880
DB 2829 GTTCCACCGCTGCTGGTTCATGTCATCTTTTATCTGTCATTTTGTACGAGCATCTCAGAA 2888
QY 881 MetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGlnPheTrpValIle 900
DB 2889 ATGAGATGGAGTGTGTGGAAATTTGATGACCTGGTGGAGGAATGAGCAGTTCTGGGTCAAT 2948
QY 901 GlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAlaGly 920

QY 190 -----LysArgIleHisProTyrProValSerGluProGly-----SerAla 203
DB 681 GGGGAAAGAGGATACATCCCTTCTTATGCGGATCCAGCTTACCTGTGCAACCCAGG 740
QY 204 LysTrpAspGluLysGlu-----ValSerTrpLysGlu 215
DB 741 TCTATGAGCCATCCAGGATCTTCTGCTATGCGTATGGTATGTTGCTTGGAGGAA 800
QY 216 ArgMetAspTrpLysSerLysGln-----GlyIleLeuGlyGly 229
DB 801 CGGATGGAGAATTGGAAGCAGACAGAGAGGATGCCACGCGGGAATGATGGTGT 860
QY 230 GlyAlaAspProGluAspMetAspAlaValAlaLeuAsnAspGluAlaArgGlnPro 249
DB 861 GGTGATGATGGTGA-----GATGCTGATCTACCACCTAATGATGAAGCAAGCAACAA 914
QY 250 LeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleVal 269
DB 915 CTGTCCAGGAAATCCACTTCCATCAAGCCAGATTATCCATATAGATGATATCAT 974
QY 270 ValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAsp 289
DB 975 ATTCCGCTTGTGGTGTCTTCTCCACTACCGAGTGATGCATCCGGTGAATGAT 1034
QY 290 AlaIleGlyLeuTrpLeuValSerIleLeuCysGluIleTrpPheAlaIleSerTrpIle 309
DB 1035 GCATTTGCTTTGTGCTCATCTCTTATCTGTGAATCTGGTTTGGCATGTCTTGGATT 1094
QY 310 LeuAspGlnPheProLysTrpPheProIleAspArgGluThrTrpLeuAspArgLeuSer 329
DB 1095 CTGTGATCAATCCCAAGTGGTCTTCTATTTGAGAGAGACTTACTAGACCGGTGTCA 1154
QY 330 LeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSer 349
DB 1155 CTGAGGTTTCGACAAGGAGCCAGCCATCTCAACTTGCTCCAAATGATTTCTTGTCTAGT 1214
QY 350 ThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeu 369
DB 1215 ACGGTGTGATCCCTTAAGAGAACCTCTTGTGTGCACAACTACTGTCTATCTATCTCTT 1274
QY 370 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 389
DB 1275 TCGGTGGATATCTCTGTGATNAGGTTTCTGTCTATGTTCTGATGATGGTGTCTGCAATG 1334
QY 390 LeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCys 409
DB 1335 CTAACGTTTGAAGCATTTATCGAAACATCTGAATTTGCCAAGAAATGGTTCCTTTCTGTC 1394
QY 410 LysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyr 429
DB 1395 AAACGGTACAAATTTGAACCTCGCGCTCCAGAGTGTGTACTTCCACAGAGATAGACTAC 1454
QY 430 LeuLysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArgGluTyr 449
DB 1455 TTGAAGAACAAGGTGGCAGCAACTTTGTTAGGAGAGAGAGCAATGAAGAGAGAT 1514
QY 450 GluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGlu 469
DB 1515 GAGGAATTCAGGTGAGAATCAATGCTTAGTTGCCAAGCCAGAAAGTTCCTGAAGAA 1574
QY 470 GlyTrpIleMetLysAspGlyThrProTrpProGlyLysAsnThrArgAspHisProGly 489
DB 1575 GGATGGACAATGCAAGATGGAACCCCTCGCGCTCCAGAGTGTGTCTTGGTATCATCCTGGA 1634
QY 490 MetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuPro 509
DB 1635 ATGATTTCAGGTCTCTCTTGGCCAAAGCGGAGCGCTTGACTGTGAGGAAATGAATGCCA 1694
QY 510 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly 529
DB 1695 CGATTTGTTTATGTTCTTAGAGAGAAACGACCGGCTATACCATCATGAAGAGCTGT 1754
QY 530 AlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsn 549

DB 1755 GCTATGAATGATTTGGTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAAC 1814
QY 550 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeu 569
DB 1815 TTGGATTGTGATCACTACATCAACACAGCAAGGCTATTAAGGAAGCAATGTGTTTATG 1874
QY 570 MetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGly 589
DB 1875 ATGGACCTTTTACTAGGAAGAAGGTTTGTCTATGTACAGTTCCCTCAAGATTTGATGGG 1934
QY 590 IleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLys 609
DB 1935 ATTGATCCCATGACCGGATATGCTAACCGGAATGTGTCTTTTGTATCAACATGAAA 1994
QY 610 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGln 629
DB 1995 GGTTTGGATGTAATCAGGTCCTAATTTATCTTGGTACTGATGTGTTATTAGAGGCAG 2054
QY 630 AlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysAsp 649
DB 2055 GCATTTATGTTTATGATGCCCCCAAA---ACAAAGAGCCACCATCAAGGACTTGAAC 2111
QY 650 -----CysCysProCysPheGlyArgLysLysArgLysHisAla 662
DB 2112 TGCTGGCCAAAGTGTGCTTTTGTGCTTGTGCTTGGCAATAGGAACAAAGAAAGACT 2171
QY 663 LysAspGlyLeuProGlu----- 668
DB 2172 ACCAAACCCAAACAGAGAGAAAAGTTATTTTCAAGAAAGAGAGAACCAATCC 2231
QY 669 -----GlyThrAlaAspMetGlyValAspSerAspLysGlu 680
DB 2232 CCTGCATATGCTCTTGGTGAATTTGACGAAGCTGCTCCAGGAGCTGAGAAAGAGGCC 2291
QY 681 MetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThr 700
DB 2292 GGTATTTGAAATCAACAAAAATTAGAAAGAAATTTGGCCAAATCTTCTGTTTGTGTACA 2351
QY 701 SerThrLeuMetGluGluGlyValProProSerSerProAlaAlaLeuLeuLys 720
DB 2352 TCCACACTTCTCGAGATGTTGGAACTTGAAGAGTGCAGTCTCTGCTCTCTTTTGAAA 2411
QY 721 GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeu 740
DB 2412 GAAGCTATACATGTCAATAGTTGTGTTATGAACAGACAGACTGGGGAAGAGAGATT 2471
QY 741 GlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg 760
DB 2472 GGCTGGATCTATGATCAGTTACAGAAGATTTCTAACTGGTTTCAAGATGCATTTGCAT 2531
QY 761 GlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIle 780
DB 2532 GGTGGCGGTCAATTTACTGCATACCTTAACCGGTTGCATTTCAAGGTTCTGCACCTCTG 2591
QY 781 AsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePhe 800
DB 2592 AATCTTTTCAGATCGTCTTCCACAGGCTCTCGGTGGGCTCTTGGTCTTATTGAGATCTTC 2651
QY 801 PheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGlu 820
DB 2652 TTCAGCAATCATTCCTCTTGTGTATGGGTAT---GGTGGCGGTCTCGAAATTTTGGAA 2708
QY 821 ArgPheAlaTyrIleAsnThrIleTyrProPheThrSerLeuProLeuLeuAlaTyr 840
DB 2709 AGATTTTCTACATCAACTCCATCGTATCTCTGGACATCTATTCCTCTTGGCTTAC 2768
QY 841 CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr 860
DB 2769 TGTACATTCGCTGCACTCTGTTTATTGACAGGAAATTTTATCACTCCAGAGCTGAATAAT 2828
QY 861 PheAlaSerLeuPheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGlu 880

Db 2829 GTTCCAGCCTGTGGTTTCATGTCACCTTTTATCTGCATCTTTTGTCTACGAGCATCTAGAA 2888
 Qy 881 MetArgTrpSerGlyValSerIleGluGluTrpTrpArgHsnGluGlnPheTrpValIle 900
 Db 2889 ATGAGATGGAGTGGTGTGGAAATGATGACTGGTGGAGGAATGACGAGTCTTGGGTCAAT 2948
 Qy 901 GlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAlaGly 920
 Db 2949 GGAGGTGTCTCCTCACACCTCTTGTCTGTCTCAGGAGCTTCTCAAGGTCATAGTGGT 3008
 Qy 921 IleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPheAla 940
 Db 3009 GTTCATCAAGCTTCCCGTGACATCAAG-----GGTGGAGATGATGAGGAGTCTCTCA 3062
 Qy 941 GluLeuTrpAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeuLeuIle 960
 Db 3063 GAGCTATATACATTCAAATGGAGTACCTTATGTATACCTCTTACACCTTGTCTTCTATTG 3122
 Qy 961 AsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTrpGlnSerTrp 980
 Db 3123 AACTTCATTGGTGTGGTCTGGCTTCAATCGGATCAATAACGGATATGATGATG 3182
 Qy 981 GlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTrpProphe 1000
 Db 3183 GGCCCCCTCTTTGGGAAGCTATTCTTTGCAATTTGGGTGATTGTCATCTTTATCCCTTT 3242
 Qy 1001 LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValValIleTrpSerIle 1020
 Db 3243 CTCAAAGGTTTGGTGGAGGCAAAACAGGACACCAACGATTTGTCATCGTCTGGTCCATT 3302
 Qy 1021 LeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThr 1040
 Db 3303 CTGCTGGCTTCATCT 3362
 Qy 1041 LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
 Db 3363 GATGTCGGCTTCTTGAGGAGTGTGGTTTGGATTGC 3398
 RESULT 7
 AAZ99503
 ID AAZ99503 standard; DNA; 3969 BP.
 AC AAZ99503;
 DT 03-JUL-2000 (first entry)
 XX DNA encoding a maize cellulose synthase.
 DE
 KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
 KW transgenic plant; plant breeding marker; ss.
 OS Zea mays.
 XX
 FH Key Location/Qualifiers
 CDS 144..3398
 FT /*tag= a
 FT /product= "cellulose synthase"
 FT /note= "no termination codon given"
 XX
 FN WO200009706-A2.
 XX
 PD 24-FEB-2000.
 XX
 PF 16-AUG-1999; 99WO-US018760.
 XX
 PR 17-AUG-1998; 98US-0096822P.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
 XX WPI; 2000-224343/19.
 DR P-PSDB; AAY84111.

XX New genes which encode maize cellulose synthase polypeptides in plants
 PT useful for modulating the expression of cellulose synthase in plants and
 PT to produce transgenic plants expressing the novel protein.
 XX
 PS Claim 1; Page 113-118; 119pp; English.
 XX
 CC The present sequence encodes a maize cellulose synthase polypeptide. The
 CC cellulose synthase can be used for the improvement of stalk quality for
 CC improved stand or silage. It also provides an increased concentration of
 CC cellulose in the pericarp, hardening the kernel and improving its
 CC handling ability. The sequences are used to produce transgenic plants and
 CC seeds expressing the cellulose synthase. The polynucleotide is used for
 CC modulating, preferably increasing, the level of the synthase in a plant
 CC cell. The plants are preferably monocots. The polynucleotide is also used
 CC as a probe or primer in the detection quantitation or isolation of gene
 CC transcripts. The probes are useful in detecting deficiencies in the level
 CC of mRNA in screenings for desired transgenic plant, for detecting or
 CC mutations in the gene, for monitoring upregulation of expression or
 CC changes in enzyme activity in screening assays of compounds, for
 CC detection of any number of allelic variants of the gene, or for use as
 CC molecular markers in plant breeding programs. The isolated nucleic acids
 CC of the present invention can also be used for recombinant expression of
 CC their encoded polypeptides or for use as immunogens in the preparation
 CC and/or screening of antibodies. The proteins can be employed in assays
 CC for enzyme agonists or antagonists of enzyme function or for use of
 CC immunogens or antigens to obtain antibodies specifically immunoreactive
 CC with a protein
 XX
 SQ Sequence 3969 BP; 1022 A; 838 C; 1018 G; 1091 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 3969
 Score: 4073.50 Matches: 757
 Percent Similarity: 79.68% Conservative: 129
 Best Local Similarity: 68.08% Mismatches: 139
 Query Match: 72.12% Indels: 87
 DB: 3 Gaps: 15
 US-10-627-132-30 (1-1052) x AAZ99503 (1-3969)
 Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
 Db 144 ATGAGGCGAGCGCGCGCTGTGTGGCGGTCTCCACAAACGCGACGAGCTGTGTCATC 203
 Qy 21 Arg-----GlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGlu 38
 Db 204 CGCGCGACGCGCATCCCGGCGGAGCCCGCGGGAGCAGAACGCGGCGGTGTGCGAG 263
 Qy 39 IleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGlu 58
 Db 264 ATTTGCGCGACGACGCTGCGCTTGCCTTCCCGCGGGGACCCCTTCGTGGCGTGCACGAG 323
 Qy 59 CysGlyPheProValCysArgProCysTrpGluArgGluArgGluGlyThrGlnAsn 78
 Db 324 TGGCCCTTCCCGCTCTGCGGGACTGCTACGAATACGAGCGCGGGAGGCGACGAGAAC 383
 Qy 79 CysProGlnCysLysThrArgTrpLysArgLeuLysGlySerProArgValAlaGlyAsp 98
 Db 384 TGCCCCCAGTGCAAGACTCGATACAAAGCGCTCAAGGGCTGCCCAACGTGTGACCGGTGAC 443
 Qy 99 AspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGln 118
 Db 444 GAGGAGGAGGACGCGCTCGATGACCTGGACACAGATTCACTGGGACGGCCATGACTCG 503
 Qy 119 GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
 Db 504 CAG-----TCTGTGGCGGAGTCCATCTCTTACGCG 533
 Qy 139 ArgMetSerTrpGlyArgGlyProAspAspGlyAsnAsnThrProGlnIlePro 158
 Db 534 CACATGAGCTACGCGCGTGG-----GGTGACCTTAATGGCGGCCCAAGCTTTC 584


```
Db 2709 AGATTTTCTACATCAATCCATCGTGTATCTTGGACATCTATTCCCTCTTGGCTTAC 2768
Qy 841 CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr 860
Db 2769 TGTACATTTGGCTGCTGCTGTTATTGACAGGAAATTTATCACTCCAGAGCTGAATAAT 2828
Qy 861 PheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGlu 880
Db 2829 GTTGCCAGCTGTGGTTCATGTCACTTTTATCTGCAATTTTGTACGACATCTAGAA 2888
Qy 881 MetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGlnPheTrpValIle 900
Db 2889 ATGAGATGGAGTGTGTGGAAATTTGATGACTGTGTGGAGGAATGAGCAGTTCTGGTCAAT 2948
Qy 901 GlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAlaGly 920
Db 2949 GGAGGTGTCTCCACACCTCTTTGCTGTGTTCAGGGACTTCTCAAGGTTCATAGCTGGT 3008
Qy 921 IleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPheAla 940
Db 3009 GTTGATCAAGCTTCACCGTGACATCAAG-----GGTGAGATGATGAGGATTTCTCA 3062
Qy 941 GluLeuTrpAlaPheLysTrpThrThrLeuLeuLeuProProThrThrLeuLeuIle 960
Db 3063 GAGCTATATACATTTCAATGGACTACCTTATTGATACCTCTTACACCTTGTCTTATTG 3122
Qy 961 AsnValIleGlyValValAlaGlyLysSerAspAlaIleAsnGlyTrpGlnSerTrp 980
Db 3123 AACTTTCATTTGGTGTGGTGGCTGGCGTTTCAATCCGATCAATAACCGATATGAGTCATGG 3182
Qy 981 GlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTrpProPhe 1000
Db 3183 GGCCCTCTTTGGGAAGCTATTTCTTGCAATTTGGGTGATGTGTCATCTTATCCCTTT 3242
Qy 1001 LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIleTrpSerIle 1020
Db 3243 CTCAAGGTTTGGTGGGAAGCAAAACAGGACACCAACGATTTGTCATCGTCTGTGTCATT 3302
Qy 1021 LeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThr 1040
Db 3303 CTGCTGGCTTCAATCTTCTCGCTCTCTTTGGGTTCGATTTGATTCCTTTCTTGGCAAGAT 3362
Qy 1041 LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3363 GATGTCCTGCTTCTTGAGGAGTGTGGTTTGATTCG 3398

RESULT 8
ID AAZ99491 standard; DNA; 3568 BP.
XX AC AAZ99491;
XX DT 03-JUL-2000 (first entry)
XX DE DNA encoding a maize cellulose synthase.
XX KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
XX KW transgenic plant; plant breeding marker; ss.
XX OS Zea mays.
XX FH Key Location/Qualifiers
XX CDS 63..3236
XX FT /*tag= a
XX FT /product= "cellulose synthase"
XX FT /note= "no termination codon given"
XX PN WO200009706-A2.
XX XX
XX PD 24-FEB-2000.
XX PF 16-AUG-1999; 99WO-US018760.
XX XX
```

```
PR 17-AUG-1998; 98US-0096822P.
XX (PION-) PIONEER HI-BRED INT INC.
XX PA Dhuga KS, Helentjaris TG, Bowen BA, Wang X;
XX PI P-PSDB; AAY84107.
XX DR WPI; 2000-224343/19.
XX DR P-PSDB; AAY84107.
XX PT New genes which encode maize cellulose synthase polypeptides in plants
XX PT useful for modulating the expression of cellulose synthase in plants and
XX PT to produce transgenic plants expressing the novel protein.
XX PS Claim 1; Page 82-87; 119pp; English.
XX CC The present sequence encodes a maize cellulose synthase polypeptide. The
XX CC cellulose synthase can be used for the improvement of stalk quality for
XX CC improved stand or silage. It also provides an increased concentration of
XX CC cellulose in the pericarp, hardening the kernel and improving its
XX CC handling ability. The sequences are used to produce transgenic plants and
XX CC seeds expressing the cellulose synthase. The polynucleotide is used for
XX CC modulating, preferably increasing, the level of the synthase in a plant
XX CC cell. The plants are preferably monocots. The polynucleotide is also used
XX CC as a probe or primer in the detection quantitation or isolation of gene
XX CC transcripts. The probes are useful in detecting deficiencies in the level
XX CC of mRNA in screenings for desired transgenic plant, for detecting or
XX CC mutations in the gene, for monitoring upregulation of expression or
XX CC changes in enzyme activity in screening assays of compounds, for
XX CC detection of any number of allelic variants of the gene, or for use as
XX CC molecular markers in plant breeding programs. The isolated nucleic acids
XX CC of the present invention can also be used for recombinant expression of
XX CC their encoded polypeptides or for use as immunogens in the preparation
XX CC and/or screening of antibodies. The proteins can be employed in assays
XX CC for enzyme agonists or antagonists of enzyme function or for use of
XX CC immunogens or antigens to obtain antibodies specifically immunoreactive
XX CC with a protein
XX SQ Sequence 3568 BP; 879 A; 825 C; 941 G; 922 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 0 Length: 3568
Score: 4005.00 Matches: 743
Percent Similarity: 79.04% Conservative: 128
Best Local Similarity: 67.42% Mismatches: 141
Query Match: 70.91% Indels: 90
DB: 3 Gaps: 16

US-10-627-132-30 (1-1052) x AAZ99491 (1-3568)
Qy 14 ArgAsnGluLeuValLeuIleArg-----GlyHisGluAspProLysProLeuArgAla 31
Db 12 CGTCCGAGTCTGTCGTCTCCGCGGATGGCGAGCCAGGCCGAGCCCATGCCAG 71
Qy 32 LeuSerGlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAsp 51
Db 72 CGGAACGCCAGGTGTGCCAGATTTCGGCGACACGACGTGGCGCGCAACCCCGCGGGAG 131
Qy 52 LeuPheValAlaCysAsnGluCysGlyPheProValCysArgProCysTrpGluTrpGlu 71
Db 132 CTTTTCGTGGCTTCAACGAGTGGCTTCCCCATCTGCCGGAGCTGCTAGATGACGAG 191
Qy 72 ArgArgGluGlyThrGlnAsnCysProGlnCysLysThrArgTrpLysArgLeuLysGly 91
Db 192 CGCCCGAGGGGCGCAGACTGCCCCAGTGCAGAGCCCGCTTCAAGCGCTTCAAGGGG 251
Qy 92 SerProArgValAlaGlyAspAspGluGlnAspIleAspAspLeuHisGluPhe 111
Db 252 TGC CGCGCGCTGCGCGGGGACGAGGAGGAGCGCGCTCGACGACCTTGGAGAACGAGTTC 311
Qy 112 AsnIleAspAspGluAsnGlnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIle 131
Db 312 AACTGGAGCGCAAGACGACGACTCCAGTACCTC----- 344
```


Db 569 GAGCGCGCGCCACAGAGAC-----GACCCGAGTACGTCGCCGAGTCC 613
Qy 135 MetLeuHisGlyArgMetSerTyrGlyArgGlyProAspAspGlyAspGly---AsnAsn 153
Db 614 ATGCTCAGGCGCAGATAGCTACGCGCGCGCGCGAGCGCCGCCGCTTCAGCC 673
Qy 154 ThrProGlnIleProProIleThrGlySerArgSerValProValSerGlyGlu--- 172
Db 674 GTCCCAACAGTCGCGCTCCTCACCACCGCCAGATGGTTGATGATCCCGCGGAGCAG 733
Qy 173 -----PhePro-----IleThrAsnGlyTyrGlyHisGlyGluValSerSerSer 187
Db 734 CACGCGCTCGTCCCTCTACATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 778
Qy 188 LeuHisLysArgIleHisProTyrProValSerGluProGly-----SerAla 203
Db 779 -----AAGAGGATCCACCGCTCCTTCGCGAGATCCCACTTCAGTGCAACCGAGA 832
Qy 204 LysTrpAspGluLysGlyGlu-----valserTrpLysGlu 215
Db 833 TCCATGAGCCCGCTCCAAAGGATCTGCCCGCTACCGGATATGGCAGCGTGGCGTGGAGGAG 892
Qy 216 ArgMetAspAspTrpLysSerLysGlnGlyIleLeu-----GlyGlyGly 230
Db 893 AGATGGAGGCTGGAGCAGAGCAGGAGCGCCCTGCAGCATGTTCAGGAGCGAGGCTGGC 952
Qy 231 AlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaAargGlnProLeu 250
Db 953 GGTGATTGGGATGGCGACGATGCAGATCTGCCACTAATGGATGAAGTAGGCGCCATTG 1012
Qy 251 SerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleValVal 270
Db 1013 TCCAGAAAGTCCCTATATATCAAGCCGAATTAATCCCTACAGGATGATTATCGTTATC 1072
Qy 271 ArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAspAla 290
Db 1073 CGTTTGGTGGTTTGGGTTCTCTCCACATCCAGGATGATGCATCCCGCGGAAAGATGCA 1132
Qy 291 IleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIleLeu 310
Db 1133 TTTGCATTGTGGCTCATATCTGTAATCTGTGAATCTGTTTGGATGTCTCGGATTCCT 1192
Qy 311 AspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSerLeu 330
Db 1193 GATCAGTTCCCAAGTGGCTTCCATCGAGAGAGAGACTTACCTGGACCGTTTGTCTACTA 1252
Qy 331 ArgTyrGluArgGluGluProSerLeuLeuSerAlaValAspPheValSerThr 350
Db 1253 AGTTTGAACAGGAAGGTCAACCCCTCTCAGCTTGTCTCCATCGACTTCTTTGTGATGACG 1312
Qy 351 ValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAla 370
Db 1313 GTTGATCCCAACAGGAACCTCCCTTGTGTACAGCGAAGACTGTCTTCCATCTCTTCT 1372
Qy 371 ValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLeu 390
Db 1373 GTGATTAATCCGTTGAGAAGTCTCTCTGCTATGTTCTTGATGATGGTGTCTGCAATGCTT 1432
Qy 391 ThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTyrValProPheCysLys 410
Db 1433 ACGTTTGAACATGTCTGAAACATCTGAATTTGCAAGAATGGGTTCTTTCAGCAAA 1492
Qy 411 LysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrIleu 430
Db 1493 AAGTTTAATAATCAGGCTCGTCTCTGAGTGGTACTTCCACAGAGATAGACTACTCTG 1552
Qy 431 LysAspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGluTyrGlu 450
Db 1553 AAAGACAGGTTGCTGCTTCAATTTGTTAGGAGAGAGGCGGCGATCAAGAGAGATACGAG 1612
Qy 451 GluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGluGly 470

Db 1613 GAATTCAGGTAAGGATCAATGCCTTGGTGTGCAAAAGCCCAAAAGGTTCTCTGAGGAAGA 1672
Qy 471 TrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGlyMet 490
Db 1673 TGGACAAATGCAAGATGGAAGCCCTGGCTGGAAACAACGTCACGCGATCATCTCTGAATG 1732
Qy 491 IleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuProArg 510
Db 1733 ATTCAAGTATTCTTGGCCAAAGTGGCGGTGCTGATGTGGAAGGAAATGAGTTGCTCGC 1792
Qy 511 LeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysValAlaGlyAla 530
Db 1793 CTGGTTTATGCTCTAGAGAAAGAGGCCAGGCTATTAACCATCAACAAGAGGCTGGTGCC 1852
Qy 531 MetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsnLeu 550
Db 1853 ATGAATGCACCTGGTCCGCTGCTCTGCTCTTATCAATGTGTCATACCTATTGAACTTG 1912
Qy 551 AspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeuMet 570
Db 1913 GACTGTGATCACTACATCAACAATAGCAAGGCCATAAAAGAGGCTATGTGTTTCATGATG 1972
Qy 571 AspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIle 590
Db 1973 GATCCTTGGTGGGGAAGAAAGTGTGCTATGTACAGTTCCCTCAGAGGTTTGATGCTATT 2032
Qy 591 AspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLysGly 610
Db 2033 GACAAAATGATCGATACGCTTAACAGGAACCTGTCTCTTTTGTGACATCAACATGAAGGT 2092
Qy 611 LeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgGlnAla 630
Db 2093 TTGACGCTATTCAAGGACCCATTTATGTGGGTACTGGATGTGTTTTCAGACGGCAGCA 2152
Qy 631 LeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysAspCys 650
Db 2153 CTGTATGTTATGATGCTCTCTTAA--ACGAAGAGGCCCATCAAGAACCTTGCAACTGC 2209
Qy 651 CysPro-----CysPheGlyArgLysLysArgLys----- 660
Db 2210 TGGCCCAAGTGGTCCCTCTCTGTCTGTGTCAGCAGGAGCAAGAATAAAAGAGACTACA 2269
Qy 660 ----- 660
Db 2270 AAACCAAGACGAGAGAAAGAAAGATTTATTTTCAAGAAAGCAGAAACCCATCTCCT 2329
Qy 661 ---HisAlaLysAspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAspLys 679
Db 2330 GCATATGCTTTGGGTGAAATTTGATGAAGGTCTCCA-----GGTGTGATATCGAAG 2383
Qy 680 GluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheVal 699
Db 2384 GCCGGAATCGTAAATCAACAGAAACTAGAGAAGAAATTTGGGCGAGTCTTCTGTTTCTG 2443
Qy 700 ThrSerThrLeuMetGluGluGlyValProProSerSerSerProAlaAlaLeuLeu 719
Db 2444 GCATCAACACTCTTGTGAGACGGAGGACCCCTGAAGAGCGCAAGTCCAGCTTCTCTCTG 2503
Qy 720 LysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGlu 739
Db 2504 AAGAACTATATACATGTTATCAGCTGGCGCTACCAAGCAAGACCCGACTGGGGAAGAG 2563
Qy 740 LeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCys 759
Db 2564 ATTGGCTGGATTACGATCGATCAGAGAGGATATCTTGATGGATTTAAGATGCACCTGC 2623
Qy 760 ArgGlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaPro 779
Db 2624 CATGGCTGGCGGTCTATTTACTGTCATCCGAAGGGCGCTGCATTCAAAGGTTCTCGCGCT 2683
Qy 780 IleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIle 799
Db 2684 CTGAACCTTTCCGACCGCTTTCACCGTCTCTCGCTGGGCGCTTGGGTCCGTCGAAAT 2743

```
QY 800 PhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeu 819
Db 2744 TTCTTTCAGCAAGCACTGCCCACTTTGGTACGGATAC-----GGCGCGGGCTAAAAATTCCTG 2800
QY 820 GluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAla 839
Db 2801 GAAAGGTTTCTTATATCAACTCATCGTTTATCCCTGGAGCTCCATTCCTCTCTGGCT 2860
QY 840 TyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSer 859
Db 2861 TACTGTACCTTGGCTGCATCTGCCTCTCGGGGAAGTTTATCACACGAGCTTACC 2920
QY 860 ThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeu 879
Db 2921 AATGTCGCCAGTATCTGGTTTCATGGCACTTTTTCATCTGCATCTCCGTCGACCGCATCTG 2980
QY 880 GluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGlnPheTrpVal 899
Db 2981 GAAATGAGGTGGAGTGGCGTGGCCATCGACACTGGTGGAGGAACGAGCAGTTCCTGGGTC 3040
QY 900 IleGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAla 919
Db 3041 ATCGAGCGCTTCGGCGCATCTGTTCGGGTGTTCAGGGCTGCTGAAGTGTTCGCC 3100
QY 920 GlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPhe 939
Db 3101 GGCATCGACAGCACTTCACCGTGACGTGCAAGGCC---GGGGAC---GACGAGGAGTTC 3154
QY 940 AlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeuLeuIle 959
Db 3155 TCGAGGTGTACAGTTCAAAGTGGACCACTCTGTGTATACCCCGACACGCTCTCTCTG 3214
QY 960 IleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSer 979
Db 3215 CTGAACCTTCATCGGGGTGGTGGCGGATCTCGAACGCGATCAACAACGGGTACGAGTCG 3274
QY 980 TrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrPro 999
Db 3275 TGGGGCCCCCTGTTCGGGAAGCTTCTTCGCCCTTCTGGGTGATCGTCCACCTGTACCCG 3334
QY 1000 PheLeuLysGlyLeuMetGlyValGlnAsnArgThrProThrValValIleTrpSer 1019
Db 3335 TTCTCAAGGTCGTGGTGGGAGGAGCAACAGGACGCGACATCGTCATCGTCTGGTCC 3394
QY 1020 IleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArg 1039
Db 3395 ATCTGTCTGGCTCGATCTTCTCGCTCTGTGGGTGGCGTGGCGTTCCTCGCCAAG 3454
QY 1040 ThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3455 AGCAACGGCCGCTCTCGGAGGAGTGTGGCCTGGACTGC 3493
RESULT 10
AAZ99509
ID AAZ99509 standard; DNA; 3813 BP.
XX
AC AAZ99509;
XX
DT 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT 215..3493
FT CDS /*tag= a
FT FT /product= "cellulose synthase"
FT FT /note= "no termination codon given"
FT
```

```
XX WO200009706-A2.
PN
XX 24-FEB-2000.
PD
XX
XX 16-AUG-1999; 99WO-US018760.
PF
XX 17-AUG-1998; 98US-0096822P.
PR
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX WPI; 2000-224343/19.
XX P-PSDB; AAY84113.
XX
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants and
PT to produce transgenic plants expressing the novel protein.
XX
PS Claim 1; Page 129-134; 119pp; English.
XX
CC The present sequence encodes a maize cellulose synthase polypeptide. The
CC cellulose synthase can be used for the improvement of stalk quality for
CC improved stand or silage. It also provides an increased concentration of
CC cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants and
CC seeds expressing the cellulose synthase. The polynucleotide is used for
CC modulating, preferably increasing, the level of the synthase in a plant
CC cell. The plants are preferably monocots. The polynucleotide is also used
CC as a probe or primer in the detection quantitation or isolation of gene
CC transcripts. The probes are useful in detecting deficiencies in the level
CC of mRNA in screenings for desired transgenic plant, for detecting or
CC mutations in the gene, for monitoring upregulation of expression or
CC changes in enzyme activity in screening assays of compounds, for
CC detection of any number of allelic variants of the gene, or for use as
CC molecular markers in plant breeding programs. The isolated nucleic acids
CC of the present invention can also be used for recombinant expression of
CC their encoded polypeptides or for use as immunogens in the preparation
CC and/or screening of antibodies. The proteins can be employed in assays
CC for enzyme agonists or antagonists of enzyme function or for use of
CC immunogens or antigens to obtain antibodies specifically immunoreactive
CC with a protein
XX
SQ Sequence 3813 BP; 873 A; 967 C; 1111 G; 862 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 0 Length: 3813
Score: 3974.50 Matches: 752
Percent Similarity: 78.53% Conservative: 122
Best Local Similarity: 67.57% Mismatches: 158
Query Match: 70.37% Indels: 81
DB: 3 Gaps: 17
XX
US-10-627-132-30 (1-1052) x AAZ99509 (1-3813)
QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
Db 215 ATGGAGGCTAGCGGGGCTGTGTGGCGGCTCGCATTAACCGAAGACGAGCTGTGTGTATC 274
QY 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnVal----- 36
Db 275 CGC-----CGCGACCGGAGTCCGGAGCCCGCGGCGCGCGCGCGCGCGGAG 328
QY 37 -----CysGluIleCysGlyAspGlnValGlyLeuThrValAspGlyAspLeuPheVal 54
Db 329 CGCGCGTGCAGATATGCGCGAGGAGTCCGGGTGGGCTTCGACGGGAGCCCTTCGTG 388
QY 55 AlaCysAsnGluCysGlyPheProValCysArgProCysTyrGluTyrGluArgArgGlu 74
Db 389 GCGTGCAACGAGTGGCCCTTCCCGTCTGCCCGCGCTCTAGAGTACGAGGCGCGCGAG 448
QY 75 GlyThrGlnAsnCysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArg 94
```

Db 449 GCTCGCAAGCGTGGCCGCGAGTGGAGGACCCGCTACAGCGCTCAAGGGCTGCCGCGG 508
Qy 95 ValAlaGlyAspAspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAsp 114
Db 509 GTGGCCGCGCAGAGGAGGAGCGCGTGGAGCGACCTGGAGGGCGAGTTGGCCCTGCAG 568
Qy 115 AspGluAsnGlnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAla 134
Db 569 GACGGCGCGCCACAGAGAC-----GACCCGACGATGACGTCGCCGAGTCC 613
Qy 135 MetLeuHisGlyArgMetSerTyrGlyArgGlyProAspAspGlyAspGly---AsnAsn 153
Db 614 ATGCTCAGGGCGCAGATGAGCTACGGCCGCGCGGCGCGCCACCCCGGCTTACGCCCC 673
Qy 154 ThrProGlnIleProProIleThrGlySerArgSerValProValSerGlyGlu--- 172
Db 674 GTCCCAACGTCGCGCTCTCCACCAACGCGCCAGATGGTTGATGACATCCCGCGGAGCAG 733
Qy 173 -----PhePro-----IleThrAsnGlyTyrGlyHisGlyGluValSerSer 187
Db 734 CACGCGCTCGTGGCTCTCTACATGAGCGCGCGCGCGCGGCG----- 778
Qy 188 LeuHisLysArgIleHisProTyrProValSerGluProGly-----SerAla 203
Db 779 -----AAGAGATCCACCGCTCTCTTCGCAGATCCCAACCTTCCAGTGCACCGAGA 832
Qy 204 LysTrpAspGluLysGlu-----ValSerTrpLysGlu 215
Db 833 TCATGGACCCGCTCCACAGATCTGGCGCTACGGATATGGCAGCGTGGCGTGGAGGAG 892
Qy 216 ArgMetAspAspTrpLysSerLysGlnGlyIleLeu-----GlyGlyGly 230
Db 893 AGAATGGAGGGCTGGAACGAGAGGAGCGCGCTGCAGCATGTGCAGGAGCGAGGGTGGC 952
Qy 231 AlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnProLeu 250
Db 953 GGTGATGGGATGGCGACGATGCAGATCTGCCACTAATGGATGAAGTAGGAGCGCCATG 1012
Qy 251 SerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleValVal 270
Db 1013 TCCAGAAAGTCCCTATATCATCAAGCGAATTAATCCCTACAGGATGATTCGTATC 1072
Qy 271 ArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAspAla 290
Db 1073 CGGTTGGTGGTTTGGGTTCTCTTCCACTACCGAGTGCATCCGCGGAAAGATGCA 1132
Qy 291 IleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIleLeu 310
Db 1133 TTGCAITGTGGCTCATATCTGTAATCTGTGAATCTGGTTGGCATGCTCGGATTCCT 1192
Qy 311 AspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSerLeu 330
Db 1193 GATCAGTTCCCAAGTGGCTTCCATCCAGAGAGAGACTTACCTGGACCGTTTGTCACTA 1252
Qy 331 ArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSerThr 350
Db 1253 AGGTTTGAAGGAGGATCAACCCCTCTCAGCTTGTCTCCAAATCGACTCTTGTTCAGTACG 1312
Qy 351 ValAspProLeuLysGluProLeuValThrAlaAsnThrValLeuSerIleLeuAla 370
Db 1313 GTTGATCCCAAGGAACCTCCCTTGGTCACAGGAACACTGTCTTCTTCCATCTTCT 1372
Qy 371 ValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyValAspMetLeu 390
Db 1373 GTGGATTATCCGTTGAGAAGTCTCTCTGCTATGTTTCTGATGATGGTGTGCAATGCTT 1432
Qy 391 ThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLys 410
Db 1433 ACGTTTGAAGCATTTGCTGAAACATCTCAATTTGCAAGAAATGGGTTCTTTTCAGCAA 1492
Qy 411 LysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeu 430

Db 1493 AAGTTTAATATATCGAGCCCTCGTGCTCTCTGAGTGGTATCTTCCAAACAGAAAGATAGACTACCTG 1552
Qy 431 LysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArgGluTyrGlu 450
Db 1553 AAAGCAAGGTTGCTGCTTCAATTTCTAGGAGAGAGGGCGATGAAGAGAGAAACGAG 1612
Qy 451 GluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGluGly 470
Db 1613 GAATCAAGGTAAGGATCAATGCTTGGTTGCAAAAGGCCAAAGGTTCTCTGAGGAAGA 1672
Qy 471 TrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGlyMet 490
Db 1673 TGGCAATGCAAGATGGAAGCCCTTGGCTGGAAACAAACAGTACGCGATCATCTCTGAAATG 1732
Qy 491 IleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuProArg 510
Db 1733 ATTCAGGTTATCTTGGCCAAAGTGGCGTCTGTATGTGGAAGGAATGAGTTGCTCGC 1792
Qy 511 LeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGlyAla 530
Db 1793 CTGGTTTATGCTTCGAGAGAAAGAGGCGCAGGTTATAACCATCACAGAAAGGCTGGTGC 1852
Qy 531 MetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsnLeu 550
Db 1853 ATGAATGCACCTGGTCCGCTCTCTGCTCTTATCAAAATGCTGCATACCTATTTGAACCTG 1912
Qy 551 AspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeuMet 570
Db 1913 GACTGTGATCATCATCAACATAGCAAGGCCATAAAGAGGCTATGTGTTTCATGATG 1972
Qy 571 AspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIle 590
Db 1973 GATCCTTTGGTGGGGAAGAAAGTGTCTATGTACAGTTTCCCTCAGAGGTTTGATGGTATT 2032
Qy 591 AspValHisAspArgTyrAlaAsnArgAsnThrValPheAspIleAsnMetLysGly 610
Db 2033 GACAAAATGATCATGATCGCTAACAGGAACCTGTCTTTTGTGACATCAACATGAAGGT 2092
Qy 611 LeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGlnAla 630
Db 2093 TTGGACGTTATTCAGGACCCATTTATGCTGGTACTGGATGTTTTCAGACGGCAGCA 2152
Qy 631 LeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysAspCys 650
Db 2153 CTGTATGTTATGATGCTCTCTAAA---ACGAAGAAGCACCACCATCAAGAACTTGCACACTGC 2209
Qy 651 CysPro-----CysPheGlyArgLysLysArgLys----- 660
Db 2210 TGGCCCAAGTGGTGCCTCTCTTGTCTGCTGCAGCAGGAACAAGAATAAAAGAAAGACTACA 2269
Qy 660 ----- 660
Db 2270 AAACCAAGACGGAGAGAGAGAAAGATATTTTCAAGAAAGCAGAGAAACCCATCTCT 2329
Qy 661 ---HisAlaLysAspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAspLys 679
Db 2330 GCATATGCTTTGGGTGAAATTTGATGAAGGCTCTCA-----GGTGTGATATCGAGAAG 2383
Qy 680 GluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheVal 699
Db 2384 GCCGGAATCGTAATCAACAGAAACTAGAGAAGAAATTTGGGCGAGTCTTCTGTTTTTCTC 2443
Qy 700 ThrSerThrLeuMetGluGluGlyValProProSerSerSerProAlaAlaLeuLeu 719
Db 2444 GCATCAACACTCTTTGAGACGGAGGACCCCTGAAGAGCGCAAGTCCAGCTTCTCTCTG 2503
Qy 720 LysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGlu 739
Db 2504 AAGGAAGCTATACATGTTATCAGCTGGCGGTACCAAGACAAGACCGACTGGGGGAAAGAG 2563
Qy 740 LeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCys 759
Db 2564 ATTGGCTGGATTTACGGATCGATCACAGAGATATCTTGACTTGGATTTAAGATGCACCTGC 2623

```
Qy 760 ArgGlyTrpArgSerValTyrCysMetProlyseArgAlaAlaPhelysGlySerAlaPro 779
Db 2624 CATGGCTGGGGCTATATTTACTGCATCCGAAGCGGCTGCATCAAAAGGTTCTGGCGCT 2683
Qy 780 IleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluile 799
Db 2684 CTGAACCTTTCCGACCGCTTTCACAGGTCCTTCGCTGGGCCCTTGGGTCCTCGAAAT 2743
Qy 800 PhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeu 819
Db 2744 TTCTTCAGCAAGCACTCCGCACCTTTGGTACGGATAC--GGCGGGCGCTAAAAATTCCTG 2800
Qy 820 GluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAla 839
Db 2801 GAAAGGTTTTCTTATATCAACTCCATCCATCTGCTTATCCCTGGACGTCATTCCTCTCGGCT 2860
Qy 840 TyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSer 859
Db 2861 TACTGTACCTTGGCTGCATCTGCTGCTCATCGGGGAAGTTTATCACACAGAGCTTACC 2920
Qy 860 ThrPheAlaSerLeuPheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeu 879
Db 2921 AATGTGCCAGTATCTGGTTTCATGCACCTTTCATCTGCATCTCCGTACCGGCATCCTG 2980
Qy 880 GluMetArgTrpSerGlyValSerIleGluTyrTrpArgAsnGluGlnPheTrpVal 899
Db 2981 GAAATGAGGTGGAGTGGCGTCCATCGACGTTGGTGGAGGAACGAGCAGTTCCTGGGTC 3040
Qy 900 IleGlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLysValLeuAla 919
Db 3041 ATCGAGGCGTTTCGGGCGCATCTGTTGCGGTGTTCAGGGCTCTGTAAGTGTTCGCC 3100
Qy 920 GlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPhe 939
Db 3101 GGCATCGACACGAGCTTCACCGTCACGTCGAGGACC--GGGGAC--GACGAGGAGTTC 3154
Qy 940 AlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeuLeuIle 959
Db 3155 TCGAGGCTGTACAGTTTCAAGTGAACCACTGCTGTATACCCCGACACGCTCTCTCTG 3214
Qy 960 IleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSer 979
Db 3215 CTGAACCTTCATCGGGTGTGGCGGGATCTCGAACGCGATCAACACGGGTACGAGTGC 3274
Qy 980 TrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrPro 999
Db 3275 TGGGGCCCGCTTTCGGGAAGCTCTTCTCGCCTTCTGGGTGATCGTCCACCTGTACCCG 3334
Qy 1000 PheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIleTrpSer 1019
Db 3335 TTCTCAAGGGTGTGGTGGGAGGAGCAAGAGGAGCGGAGCGATCGTCTGTGGTCC 3394
Qy 1020 IleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArg 1039
Db 3395 ATCTCTGCTGGCTCGATCTTCTGGTCTCTGTGGGTCCGGGTCCGACCCGTTCTCGCCAG 3454
Qy 1040 ThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3455 AGCAACGCGCGCTCTCTGGAGGAGTGTGGCTGGACTGC 3493
```

RESULT 11

AAZ99524

ID AAZ99524 standard; DNA; 3813 BP.

XX AC AAZ99524;

XX XX

DT 03-JUL-2000 (first entry)

XX DNA encoding a maize cellulose synthase.

DE Maize; cellulose synthase; stalk quality; stand; silage; cellulose;

KW transgenic plant; plant breeding marker; ss.

```
XX OS Zea mays.
XX FH Key Location/Qualifiers
XX CDS 215..3493
FT /*tag= a
FT /product= "cellulose synthase"
FT /note= "no termination codon given"
XX
```

WO200009706-A2.

XX

XX

PD 24-FEB-2000.

XX

PF 16-AUG-1999; 99WO-US018760.

XX

PR 17-AUG-1998; 98US-0096822P.

XX (PION-) PIONEER HI-BRED INT INC.

XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;

XX WPI; 2000-224343/19.

DR P-PSDB; AAY84118.

XX New genes which encode maize cellulose synthase polypeptides in plants useful for modulating the expression of cellulose synthase in plants and to produce transgenic plants expressing the novel protein.

PS Claim 1; Page 168-173; 119pp; English.

XX The present sequence encodes a maize cellulose synthase polypeptide. The cellulose synthase can be used for the improvement of stalk quality for improved stand or silage. It also provides an increased concentration of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting or mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein

XX SQ Sequence 3813 BP; 873 A; 967 C; 1111 G; 862 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	3813
Score:	3974.50	Matches:	752
Percent Similarity:	78.53%	Conservative:	122
Best Local Similarity:	67.57%	Mismatches:	158
Query Match:	70.37%	Indels:	81
DB:	3	Gaps:	17

US-10-627-132-30 (1-1052) x AAZ99524 (1-3813)

Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuile 20

Db 215 ATGGAGGCTAGCGGGGCTGGTGGCGGCTCGCATTAACCGGAACGAGCTGGTGGTGATC 274

Qy 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnVal----- 36

Db 275 CGC-----CCGACCGCGAGTCGGAGCCCGGGCGCGCGCGCGCGCGGAG 328

Qy 37 -----CysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheVal 54


```
Qy 720 LysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrrpGlyLeuGlu 739
Db 2504 AAGAGCTATACATGTTATCAGCTGGCTGCTAGAGACAGACCGAGCTGGGGAAGAG 2563
Qy 740 LeuGlyTrrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCys 759
Db 2564 ATTGGCTGGATTACGATCGATCGATCAGAGAGATATCTTGACTGGATTAAAGATGCACTGC 2623
Qy 760 ArgGlyTrrpArgSerValTyrCysMetProLysArgAlaIlePheLysGlySerAlaPro 779
Db 2624 CATGGCTGGGGCTCTATTACTGCATCCGAGCGGCTGCATTCRAAGGTTCTGCGCCT 2683
Qy 780 IleAsnLeuSerArgLeuAsnGlnValIleuArgTrrpAlaLeuGlySerValGluIle 799
Db 2684 CTGAACCTTTCCGACCGCTTCACAGGTCCTTCGCTGGGCGCTTGGGTCCGTCGAAT 2743
Qy 800 PhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrrpLeu 819
Db 2744 TTCTTCAGCAAGCACTGCCACATTGGTACGGATAC---GGCGGGGGCTAAAATTCTCTG 2800
Qy 820 GluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAla 839
Db 2801 GAAAGGTTTCTTATATCAATCCATCGTTTATCCCTGGACGTCATTCCTCTCTGGCT 2860
Qy 840 TyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSer 859
Db 2861 TACTGTACCTTGCCTGCATCTGCCTGCACGGGAGTTTATCACACAGAGCTTACC 2920
Qy 860 ThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeu 879
Db 2921 AATGTCGCCATATCTGGTTATCGCATCTTCATCTGCATCTCCGTCGCGCATCTCTG 2980
Qy 880 GluMetArgTrrpSerGlyValSerIleGluIuTrrpArgAsnGluGlnPheTrrpVal 899
Db 2981 GAAATGAGTGGAGTGGCGGTCGATCGACACTGGTGGAGGAACGAGCACTTCTGGGTC 3040
Qy 900 IleGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLysValLeuAla 919
Db 3041 ATCGAGGCGTTTCGGCGCATCTGTTCGGGTGTTCCAGGGCCCTGTAAGGTGTTCCGC 3100
Qy 920 GlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPhe 939
Db 3101 GGCAATCGACAGAGCTTCACCGTCAGCTGCAAGGCC---GGGGAC---GACGAGGAGTTC 3154
Qy 940 AlaGluLeuTyrAlaPheLysTrrpThrThrLeuLeuIleProProThrThrLeuLeuIle 959
Db 3155 TCGGAGCTGTACAGTTCAGTGGACCACTGCTGATACCCCGGACCGCTCTCTCTG 3214
Qy 960 IleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSer 979
Db 3215 CTGAACCTTCATCGGGGTGGTGGCGGATCTCGAACCGGATCAACACGGGTACGAGTCG 3274
Qy 980 TrrpGlyProLeuPheGlyLeuPhePheAlaPheTrrpValIleValHisLeuTyrPro 999
Db 3275 TGGGGCCCCCTGTTCCGGAAGCTTCTTCGGCTTCTGGGTGATCGTCCACCTGTACCCG 3334
Qy 1000 PheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIleTrrpSer 1019
Db 3335 TTCTCTCAAGGCTCTGGTGGGAGGAGCAACAGGACCGCGAGATCGTATCGTCTGTGTC 3394
Qy 1020 IleLeuLeuAlaSerIlePheSerLeuLeuTrrpValArgIleAspProPheIleValArg 1039
Db 3395 ATCTCTGGGCTCGATCTTCGCTCTGCTGGGTCCGCGGACCGCTTCTCTGCCAAG 3454
Qy 1040 ThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3455 AGCAACGGCCCGCTCTCTGGAGGAGTGGCTGGACTGC 3493
RESULT 12
AAC49550
ID AAC49550 standard; DNA; 3198 BP.
XX
```

```
AC AAC49550;
XX 18-OCT-2000 (first entry)
DT Arabidopsis thaliana DNA fragment SEQ ID NO: 61572.
XX Arabidopsis thaliana.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
PD 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 03-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
```

```
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145113P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145226P.
PR 27-JUL-1999; 99US-0145911P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149360P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149428P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150568P.

PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 0
Score: 3932.00
Percent Similarity: 79.17%
Best Local Similarity: 68.81%
Query Match: 69.62%
DB: 3

Length: 3198
Matches: 750
Conservative: 113
Mismatches: 139
Indels: 88
Gaps: 18

US-10-627-132-30 (1-1052) x AAC49550 (1-3198)

QY 22 GlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGluIleCysGly 41
DB 13 GGAGAAACCGCGGAGAGCCGATGAGACATTGTTCCGAGACTTGCAGACTCTGTAGT 72
QY 42 AspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGlyPhe 61
DB 73 GACAAATGTTGGCAAGACTGTTGATGGAGATCGTTTTGGCTTGTGATATTGTTCAATTC 132
QY 62 ProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsnCysProGln 81
```


133 CCAGTTTGTGGCGCTTCTAGAGTATGAGAGGAAAGATGGAATCAATCTTGTCTCTCAG 192
Qy
82 CysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAspAspGlu 101
Db
193 TGC AAAACCAAGAGAGCTCAAGAGGTAGTCTCTATTCCTGGTGATAAGACGAG 252
Qy
102 GluAspIle--AspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnArg 120
Db
253 GATGGCTTGTAGTGAAGGTACTGTGTGATTTCACTTACCTT----- 294
Db
121 GlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMet 140
Qy
295 -----CAGAAGGAGAAATTTTCAGAGCGGATGCTTGTGTGCATCTT 336
Db
141 SerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIlePro----- 158
Qy
337 ACTCGTGGGAGGAGGAGAAATGGGGAAACCCAGTATGATAAAGAGGTCTCTCACAA 396
Db
159 -----ProIleIleThrGlySerArgSerValProValSerGluPhe----- 173
Qy
397 CATCTTCTGTCTCACGACGACAGCAAGAT-----ACTTCAGGAGAGTTTCTGTGCGC 450
Db
174 -----ProIleThrAsnGly-----TyrGly 180
Qy
451 TCACCTGAACGCTCTCTGTATCTTCTACTATCGCTGGGGAAAGCGCTTCCCTATTCA 510
Db
181 HisGlyGluValSerSerLeuHisLysArgIleHisProTyrProValSerGluPro 200
Qy
511 ---TCAGATGTCAATCAATCAACCAATAGAGGAT-----GTGGATCTCTGTT 555
Db
201 GlySerAlaLysTyrAspGluLysGluValSerTrpLysGluArgMetAspAspTrp 220
Qy
556 GGACTCGGG-----AATGTACTTGGAAAGGAGAGGTGTATGGCTGG 597
Db
221 LysSerLysGln-----GlyLeuGly-----Gly 229
Qy
598 AABATGAAGCAAGAGAAATFACTGGTCTCTCAGCAGCGAGGCTGCTTCTGAAAGAGGT 657
Db
230 GlyAlaAspProGlu-----AspMetAspAlaAspValAla----LeuAsnAspGlu 245
Qy
658 GGAGTAGATATTGATGCCAGCACAGATATCTCTAGCAGATGAGGCTCTGCTGAATGACGAA 717
Db
246 AlaArgGlnProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArg 265
Qy
718 CGAGGCGAGCTCTGTCAAGAAAGTTTCAATTCCTTCATCCAGATCAATCTTACAGA 777
Db
266 MetValIleValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHis 285
Qy
778 ATGGTTATTATGTGCGGCTTGTATTCTTGTCTCTTCTTGTGCATTAACGATATAACAAAC 837
Db
286 ProValProAspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAla 305
Qy
838 CCAGTGC AAAATAGCCTTGTCTATGGCTGTCTGTGTATGTGAGATCTGTTGTC 897
Db
306 IleSerTrpIleLeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeu 325
Qy
898 TTAICTCTGGATTTTGGATTCAGTTTCCCAAGTGGTTTCTGTGAACCGTGAACCTACCTC 957
Db
326 AspArgLeuSerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAsp 345
Qy
958 GACAGGCTTGTCTTTAAGATATGATCGTGAAGGTGAGCCATCACAGTTAGCTGCTGTTCAC 1017
Db
346 LeuPheValSerThrValAspProLeuLysGluProProLeuValThrAlaAsnThrVal 365
Qy
1018 ATTTTCTGTGAGTACTGTGACCCCTTGAAGGAGGCCACCCCTTGTGACAGCCAAACAGTG 1077
Db
366 LeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAsp 385
Qy
1078 CTCTCTATTCTGGCTTGTGATCCAGGTGTGACAGGTGTCTGTATGTCTTCTGTATGAT 1137
Db
386 GlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrp 405
Qy
1138 GGTGCTGCTATGTTATCATTTTGAATCACTTTCAGAGAAACATCAGAGTTTGTCTCGTAAATGG 1197
Db

Qy
406 ValProPheCysLysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeu 425
Db
1198 GTACCAATTTTGCAGAAATATACCATAGAGCTCTGTCACCAAGATGGTACTTTCTGCTCG 1257
Qy
426 LysValAspTyrLeuLysAspLysValGlnProThrPheValGlnGluArgAlaMet 445
Db
1258 AAAATAGATTACTTTGAAGGATAAAGTTTCAGACATCATTTTGTCAAAGATCGTAGAGCTATG 1317
Qy
446 LysArgGluTyrGluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLys 465
Db
1318 AAGAGGGAATATGAGGAATTTAAATCCGAATCCATGCACTTGTGTCCAAAGCCCTAAAA 1377
Qy
466 ValProAlaGluGlyTrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArg 485
Db
1378 TGTCTCTGAAGAGGGTGGGTATGCAAGATGGACACCGCTGGCTGGNAATATATACAAGG 1437
Qy
486 AspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGly 505
Db
1438 GACCATCCAGGAATGATCCAGGTCTTTTAGGGCAAAATGGTGGACTTGTATGCAGAGGGC 1497
Qy
506 AsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHis 525
Db
1498 AATGAGCTCCCGGCTTGTGTATATGTTCTCGAAGAAAGCGACAGATTCAGACCCAC 1557
Qy
526 LysLysAlaGlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaPro 545
Db
1558 AAAAGGCTGGTCTATGATGACCTGGTGAGAGTTTCAGACGTTCTTACCAATGGACCT 1617
Qy
546 PheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAla 565
Db
1618 TTCACTTTGAATCTTGATGTGATCATTAATAATAACAGCAAAAGCCTTTAAGAGAAGCA 1677
Qy
566 MetCysPheLeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGln 585
Db
1678 ATGTGCTTCTGTATGAGCCAAACCTCGGGAAGCAAGTTTGTATGTTCAGTTCCCAACAA 1737
Qy
586 ArgPheAspGlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAsp 605
Db
1738 AGATTGTATGATTCGATAGACGATAGATATGCTAATCGTAATACCGTGTCTTCTTGTAT 1797
Qy
606 IleAsnMetLysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysVal 625
Db
1798 ATTAACCTTGAGAGGTTTAGATGGATTCAAGGACCTGTATATGTCGGAACCTGGATGTGT 1857
Qy
626 PheArgGlnAlaLeuTyrGlyTyrAsnProPro-----LysGlyProLysArgPro 643
Db
1858 TTCAACAGACAGCATTTATACGGTTATGAACCTTCAATAAAAGTAAACACCAAGAGCCA 1917
Qy
644 LysMetValThrCysAspCysProCysPheGlyArgLysLysArgLysHisAlaLys 663
Db
1918 AGCTTTTATCTAAGCTCTGTGGTGA-----TCAAGAAAGAAAGAAATTCCAAAGCTAAG 1971
Qy
663 ----- 663
Db
1972 AAAGAGTCGCAAAAAGAAATCAGCGAGGCATACTGACTCAACTGTTCCTGTATTCAAC 2031
Qy
664 ---AspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAspLysGluMetLeu 682
Db
2032 CTCGATCAGCATAGAAAGGAGGTTGAAGGTCGTGGTTTGTATGATGATAAAGGCGCTCTTA 2091
Qy
683 MetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThrSerThr 702
Db
2092 ATGTCCGAATGAGCTCGGAGACCGATTTGGACAGCTGCTGTTTTTGTGTTCTTACC 2151
Qy
703 LeuMetGluGluGlyValProProSerSerSerProAlaAlaLeuLeuLysGluAla 722
Db
2152 CTAATGCAAAATGGTGTCTCTCTTCAGCACTCCAGAAAAACCTTCTCAAAGAGGCT 2211
Qy
723 IleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeuGlyTrp 742
Db
2212 ATCCATGTCTATTGTGTTATGAGATAAGTCAGATTGGGAATGGAGATTGGATGG 2271
Db

Qy	743	IleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArgGlyTrp	762
Db	2272	ATCTATGGTTCTGTGCACAGAGATATTCTCAGCTGGGTTCAAAATGCATGCCCGTGGATGG	2331
Qy	763	ArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIleAsnLeu	782
Db	2332	CGATCCATTACTGCATGCCCTAAGCTTCCAGCTTTCAGGGTCTTGCTCTCTATCAATCTT	2391
Qy	783	SerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePhePheSer	802
Db	2392	TCAGATCGTCTGAACCAAGTCTCAGGTGGCTTTAGGTTTCAGTTGAGATTCTCTTCAGT	2451
Qy	803	ArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGluArgPhe	822
Db	2452	CGGCATTGTCTCTATGATGATGGTTAC--AATGGGAGGCTAAATAATCTTCAGAGGTTT	2508
Qy	823	AlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyrCysThr	842
Db	2509	GGGTATGTGAACACCACCATCTACCTTCACCTCCATTCTCTCTTCATGTATTGTACA	2568
Qy	843	LeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThrPheAla	862
Db	2569	TTGCCACCGTTGTCTCTCTCCACCAACCAAGTTATTATCTCTCAGATTAGTACATTGCA	2628
Qy	863	SerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyLysLeuGluMetArg	882
Db	2629	AGTATATGGTTTCTGTCTCTCTCTCTCCATTTCGCCACGGGTATCTAGAAATGAGG	2688
Qy	883	TrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGlnPheTrpValIleGlyGly	902
Db	2689	TGGAGTGGCGTAGGCATAGACGAATGTGGAGAAACGACGAGTTTGGGTCATTGGTGA	2748
Qy	903	ValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAlaGlyIleAsp	922
Db	2749	GTATCCGCTCATTTATTTCGCTGTGTTCAGGTATCTCTCAAGTCTCTGCGGTATTGAC	2808
Qy	923	ThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPheAlaGluLeu	942
Db	2809	ACAACTCTCACAGTTACTCAAAAGCTTCA--GATGAACGCGAGACTTGTCTGAGCTC	2865
Qy	943	TyrAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeuLeuIleIleAsnVal	962
Db	2866	TACTTGTTCAAATGGACACACACTCTGATTCGCCAACGACGCTGCTCATTTGAACCTTA	2925
Qy	963	IleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTrpGlyPro	982
Db	2926	GTGGGAGTTCTTGCAGGAGTCTCTTATGCTATCAACAGTGGATACCAATCATCGGACCA	2985
Qy	983	LeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLys	1002
Db	2986	CTCTTTGGTAAATGTGTTCTTTTGGCTCTCTGGGTGATTGTTTCACTTGTACCTTCCTCAAG	3045
Qy	1003	GlyLeuMetGlyArgGlnAsnArgThrProThrValValValIleTrpSerIleLeuLeu	1022
Db	3046	GGTTTGAATGGGTGCACAGAACCGGACTCCCTACCATTTGTGTGGTGTGCTGTCTCTTG	3105
Qy	1023	AlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThrLysGly	1042
Db	3106	GCTTCTATCTCTCTGTGTGTGGGTAGGATTGATCCCTTCTACTACCGCAGTCACTGGC	3165
Qy	1043	ProAspValArgGlnCysGlyIleAsnCys	1052
Db	3166	CCGGACATTCTGGAAATGTGAACTCAACTGT	3195

RESIT.T 13

RESULI 13
AAS16458

AA316438
ID AAS16458 standard: cDNA: 3799 BP.

XX
00401000

AC AAS16458;

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818

DT 14-FEB-2002 (first entry)

XX
11 FEB 2004 11100Z ENHLY/

Corn; ss; cellulose synthase; Cdp945; ceaA-3;
stalk quality; improved seed; silage; pericar-
pal handling ability; transgenic plant.

Zea mays.

Key : Location/Qualifiers
CDS 238..3799
/*tag= a
/product= "Cellulose_synthase"

WO200179516-A2.

25-OCT-2001.

12-APR-2001: 2001WO-US011951.

100

14-APR-2000; 2000US-0055048

(PION-) PIONEER HT-BE

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Drugga KS, Helentjaris TG;
WPI: 2002-041338/05.
P-PSDB; AAU0496.
New cellulose synthase polypeptides and polynucleotides, useful in improving stalk quality or silage, and in increasing concentration of cellulose in the pericarp, hardening the kernel for improved handling ability.

Claim 4: Page 80-85; 88pp; English.

100

synthase proteins from corn, Cdpqs45 (cesa-3) and Cqrae19 (

disclosed are a recombinant expression cassette comprising

polynucleotide (operably linked to a promoter) a host cell recombinant expression cassette and a transgenic plant com-

recombinant expression cassette and a transgenic plant comprising the recombinant expression cassette. The nucleic acid is useful

improvement of stalk quality for improved stand or silage,

increased concentration of cellulose in the pericarp, harder kernel, and thus improving its handling ability. The nuclei

also be used as probes or amplification primers in the detection, and thus improving its handling ability. The nucleic

quantification or isolation of gene tran

deficiencies in the level of mRNA, for detecting gene mutations or allelic variants for monitoring an accumulation of environmental changes

allelic variants, for monitoring up-regulation or expression of changes in enzyme activity in screening assays for site directed mutagenesis.

and in sense or antisense suppression of one or more genes in a host

cell, tissue or plant. The polypeptides may be used in assays for enzyme

Qy	74	GlulGlyThrGlnAsnCySPProGlnCysIleThrArgTyrIleArgLeuLysGlySerPro	93
Db	403	GAGGGCACACAGCGTCCCCAGTGCMAAAACAAGTACAAGGCCACAAGGGAGTGCCA	462
Qy	94	ArgValAlaGlyAspAsp---AspGluGluAspIleAspAspLeuGluHisGluPheAsn	112
Db	463	GCATCCGAGGGGAGGAGAGACGATCATGTATGTCGCGATGATGCTAGC---GACTTCAAC	519
Qy	113	Ile-----AspAspGluAsnGlnGlnArgGlnLeuGluGlyAsnMetGln	127
Db	520	TACCTGCATCTGGCAATGACGACGAGCAG-----	552
Qy	128	AsnSerGlnIleThrGluAlaMetLeuHisGlyArgMetSerTyrGlyArgGlyProAsp	147
Db	553	-----AAGATGTGTGACAGGATCGCGAGCTGGCGCATGAATCTGGGGCGACGCGGAT	606
Qy	148	AspGly-----AspGlyAsnAsn	153
Db	607	GTTGGCGCGCCCCAAGTATGACAGTGTGTGAGATCGGGCTTACCAAGTACCAAGTGTGTGAG	666
Qy	154	ThrProGln---IleProProIleIleThrGlySerArgSerValProValSerGlyGlu	172
Db	667	ATCCTCGGGATACATCCCGTCAGTCACCTAACAGCCAG-----ATTTCGGGAGAA	717
Qy	173	PheProIleThrAsnGlyTyrGlyHisGlyGluValSerSerSerLeuHisIysArgIle	192
Db	718	ATCCCTGGTGTCTCCCTGCACCATCATATGATGTCTCTACTCGGGAACATTTGGCAGCGC	777
Qy	193	HisProTyrProVal-----SerGluProGlySerAlaIlystrpAspGluLys	208
Db	778	GCCCATCTCCCTATATGATGATCATTCATCAATCCGTCGAGGGGAATCTCTGGTAGCGTT	837
Qy	209	LysGluValSerTrpLysGluArgMetAspAspTrpLysSerLysGln-----GlyIle	226
Db	838	GGGAATGTGTGCTGGAAGAGAGGGTTGATGTGCTGGAAAAATGAAGCAGGACAAGGGAACA	897
Qy	227	Leu-----GlyGlyGlyAlaAspProGluAspMet	236
Db	898	ATTCCTCATGACGAATGGCACAGCATTTGCTCCCTCTGAGGGCGGGGTGTGTGTGATAT	957
Qy	237	AspAla-----AspValAlaLeuAsnAspGluAlaArgGlnPro	249
Db	958	GATGCATCACTGATTACAACATGGAAGATGCCCTTTATTAACCATGAATCCGCCAGCCT	1017
Qy	250	LeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleVal	269
Db	1018	CTATCTAGAAAGTTCCCATCTCTCTCCCTCCAGGATAATCATACAGGATGGTCATTGTG	1077
Qy	270	ValArgLeuValValLeuAlaPheLeuArgTyrArgIleLeuHisProValProAsp	289
Db	1078	CTAGATTGATGTCTTAGCATCTTTTGCACTACCGGATCAAAATCCTGTGCGTAAAT	1137
Qy	290	AlaIleGlyLeuTrpLeuLysIleIleCysGluIleTrpPheAlaIleSerTrpIle	309
Db	1138	GCATACCCACTGTGGCTTCTATCTGTTATATGTGAGATCTGGTTGCTCTTCTCTCGATA	1197
Qy	310	LeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSer	329
Db	1198	TTGGATCATGTTCCAAAGTGGTTTCCAAATCAACCGCGAGACTTACCTTGTATAGATCGCA	1257
Qy	330	LeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSer	349
Db	1258	TTAAGTATGACCGGAGAGGTGAGCCATCTCAGTTGGCTGCTGTGTGACATTTTGTTCAGT	1317
Qy	350	ThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeu	369
Db	1318	ACTGTGCAACCAATGAGGAGCGCTCTCTGTCTCATCGCAATACCGTGCATTCATTCCTC	1377
Qy	370	AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet	389
Db	1378	GCTGTGGACTATCTCTGTGTGATAAGGTCTCTTGTCTATGTATCTGATGATGAGCTGCTATG	1437

390	LeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCys	409
1438	CTGACATTTGATGCGCACTAGCTCGAGACTTCAGAGTTTCTAGAAAAATGGGTGGCCATTTGTT	1497
410	LysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyr	429
1498	AAGAAGTACAAACATTAACCTAGAGCTCTCGAATGGTACTTCTCCACAGAAATATTGATTAC	1557
430	LeuLysAspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGluTyr	449
1558	TTGAGGACAAAGTCACCTTCATTTGTTAAAGACCGCGGCCCATGAAGAGAGAATAT	1617
450	GluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGlu	459
1618	GAAGAATTCAAATTAGGGTAATGGCTTTGTCTAAAGGCACAAAAAGTCCCTCAGGAA	1677
470	GlyTyrIleMetLysAspGlyThrProTyrProGlyAsnAsnThrArgAspHisProGly	489
1678	GGATGGATCATGCAAGATGGCAACCATGGCCAGGAAACAAATACCAAGGGACCAATCCTGG	1737
490	MetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuPro	509
1738	ATGATTCAGGTTTTCTTGGTACAGTGGTGGCTTGATATCGAGGTAAATGAGCTACCC	1797
510	ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysLysAlaGly	529
1798	CGTTTGGTCTATGTTCTCGTGAACACGCTCTGGATTCGAGCATCAACAGAAAGCTGGT	1857
530	AlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsn	549
1858	GCCATGAATGCTCTGTGCCGCTCTCAGCTGTGCTTACCAATGGACAAATACATGTTGAAT	1917
550	LeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeu	569
1918	CTTGATTGTGATCATCTACATCAACACAGTAAAGCTCTCAGGAGAGCTATGTGCTTCCTT	1977
570	MetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGly	589
1978	ATGGATCCTTAACCTAGGAAGAGTGTCTGCTATGTTTCAGTTTCCCCAGAGGTTCGATGGT	2037
590	IleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleLeuAsnMetLys	609
2038	ATTGATAGGAATGATCATGATATGCCACAGAAACACCGTGTTTTTCGATATTAACTTGAGA	2097
610	GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGln	629
2098	GGTCTTGATGGCATCCAGAACCCAGTTTATGTGGCACTGGCTGTGTTTTCAACAGAGA	2157
630	AlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysAsp	649
2158	GCTCTATATGTTATGATGCCCCCAATTAAGCMAAAGAAGGTGGTGTCTTCTGTCTCA	2211
650	CysCysProCysPheGlyArgLysLys	658
2212	---TCACTATGTGTTAGGACCCCAATTAAGCMAAAGAAGGTGGTGTCTTCTGTCTCA	2268
659	-----ArgLysHisAlaLysAspGlyLeuPro	668
2269	AAGTCACAGAAGCATGTGGACAGTCTGTGGCCATATTCAATCTTGAAGATATAGAGG	2328
669	GlyThrAlaAspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPhe	688
2329	GGAGTTGAAGCGCTGGATTGTGATGATGAGAAATCATCTTCTTATGTCTCAATGAGCTTG	2388
689	GluLysArgPheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGlyGly	708
2389	GAGAAGAGATTGGCCAAATCTGCAGCTTTTGTTCGTCCTCATCTGATGGAATATGTTGGT	2448
709	ValProProSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCys	728
2449	GTTCTCAGTCTGGACATCCAGAAATCTCTTCTGTAAGAGAGATATCCATGCTCATAGTTGT	2508
729	GlyTyrGluAspLysThrAspTrpGlyLeuGluLeuGlyTyrIleTyrGlySerIleThr	748

```
Db 2509 GGCCTACGAGACAGATTGAATGGGAACTGAGATGGGTGGATCTATGTTCTGGACG 2568
Qy 749 GluAspIleLeuThrGlyPheLysMetHisCysArgGlyTrpArgSerValTyrCysMet 768
Db 2569 GAAGATATTTCTCACTGGGTTCAGATGACACGACGAGGCTGGGGTTCGATCTACTGCATG 2628
Qy 769 ProLysArgAlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAenGln 788
Db 2629 CCTAAGCGCGCGGCTTCAAGGGATCGGCTCCCATCAATCTCTCAGACCGTCTGAACCCAG 2688
Qy 789 ValLeuArgTrpAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeu 808
Db 2689 GTGCTCCGGTGGGCTCTCGTTTCAGTGGAAATCTTTTCAGCGGATGTCCTCCCTATGG 2748
Qy 809 TyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrThr 828
Db 2749 TACGGGTAC--GGAGGAGCGCTGAAGTTCTTTGGAGAGATTCCCTACATCAACACCACC 2805
Qy 829 IleTyrProPheThrSerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeu 848
Db 2806 ATCTACCCGCTCACGCTCCCTCCGCTCTCTCATTTACTGTATCTCGCTGCCATCTGCCTG 2865
Qy 849 LeuThrGlyLysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIleAla 868
Db 2866 CTCACGGGAAGTTTCATCATCCAGAGATCAGAACTTCGCTAGTATCTGGTTCATCTCT 2925
Qy 869 LeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIle 888
Db 2926 CTCCTTCATCTCGATCTCCGACGCGGTATCTCGAGATGAGTGAGCGGCGTGGGCATC 2985
Qy 889 GluGluTrpTrpArgAsnGluGlnPheTrpValIleGlyGlyValSerAlaHisLeuPhe 908
Db 2986 GACGAGTGGTGGAGAACGAGCAGTTCTGGGTTCATCGGAGGATCTCCGCCACCTCTTC 3045
Qy 909 AlaValValGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThr 928
Db 3046 GCGGTCTTCAGGGCTCTCTCAAGGTGCTTGGCGGATCGACCACTTCACCGTCACC 3105
Qy 929 SerLysAlaThrGlyAspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThr 948
Db 3106 TCCAGGGCTCG--GATGAAGACGCGGACTTCGCGGAGCTGTACATGTTCAAGTGGACG 3162
Qy 949 ThrLeuLeuIleProProThrThrLeuLeuLeuIleLeuAsnValIleGlyValAlaGly 968
Db 3163 ACACCTTCTGATCCCGCCACCACTCTCTGATCATCAACCTGGTGGGCTGTGTTCCCGCG 3222
Qy 969 IleSerAspAlaIleAsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhe 988
Db 3223 ATCTCTACGCCATCAACAGCGGGTACCAGTCTGGGGTCCGCTCTTCGGCAAGCTCTTC 3282
Qy 989 PheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGln 1008
Db 3283 TTCGCTCTTGGTGTATGTTTCACCTGTACCGTCTCTCAAGGCTCTATGGTGGGACG 3342
Qy 1009 AsnArgThrProThrValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeu 1028
Db 3343 AACCCACCCCGACCATCTGTTGTTCTCTGGGCGATCTCTGCTGGCTCGATCTCTCTCTG 3402
Qy 1029 LeuTrpValArgIleAspProPheIleValArgThrLysGlyProAspValArgGlnCys 1048
Db 3403 CTGTGGGTTCGATCGATCCGTTCCACCAACCGCGTCACTGGCCCGGATCTCGAACGTTG 3462
Qy 1049 GlyIleAsnCys 1052
Db 3463 GGCATCACTGC 3474
```

RESULT 14

AAZ99512

ID AAZ99512 standard; DNA; 3746 BP.

XX

AC AAZ99512;

XX

```
DT 03-JUL-2000 (first entry)
XX DNA encoding a maize cellulose synthase.
XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX Zea mays.
XX Key Location/Qualifiers
XX CDS 321..3548
XX /*tag= a
XX /product= "cellulose synthase"
XX /transl_except= (pos: 1800..1802, aa: Xaa)
XX /note= "no termination codon given; Xaa is an unspecified
XX amino acid"
XX WO200009706-A2.
XX 24-FEB-2000.
XX 16-AUG-1999; 99WO-US018760.
XX 17-AUG-1998; 98US-0096822P.
XX (PION-) PIONEER HI-BRED INT INC.
XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX WPI; 2000-224343/19.
XX P-PSDB; AAY84114.
XX New genes which encode maize cellulose synthase polypeptides in plants
XX useful for modulating the expression of cellulose synthase in plants and
XX to produce transgenic plants expressing the novel protein.
XX Claim 1; Page 137-141; 119pp; English.
XX The present sequence encodes a maize cellulose synthase polypeptide. The
XX cellulose synthase can be used for the improvement of stalk quality for
XX improved stand or silage. It also provides an increased concentration of
XX cellulose in the pericarp, hardening the kernel and improving its
XX handling ability. The sequences are used to produce transgenic plants and
XX seeds expressing the cellulose synthase. The polynucleotide is used for
XX modulating, preferably increasing, the level of the synthase in a plant
XX cell. The plants are preferably monocots. The polynucleotide is also used
XX as a probe or primer in the detection quantitation or isolation of gene
XX transcripts. The probes are useful in detecting deficiencies in the level
XX of mRNA in screenings for desired transgenic plant, for detecting or
XX mutations in the gene, for monitoring upregulation of expression or
XX changes in enzyme activity in screening assays of compounds, for
XX detection of any number of allelic variants of the gene, or for use as
XX molecular markers in plant breeding programs. The isolated nucleic acids
XX of the present invention can also be used for recombinant expression of
XX their encoded polypeptides or for use as immunogens in the preparation
XX and/or screening of antibodies. The proteins can be employed in assays
XX for enzyme agonists or antagonists of enzyme function or for use of
XX immunogens or antigens to obtain antibodies specifically immunoreactive
XX with a protein
XX Sequence 3746 BP; 861 A; 976 C; 1006 G; 902 T; 0 U; 1 Other;
```

Alignment Scores:

Pred. No.:	0	Length:	3746
Score:	3912.50	Matches:	750
Percent Similarity:	78.96%	Conservative:	102
Best Local Similarity:	69.51%	Mismatches:	150
Query Match:	69.27%	Indels:	77
DB:	3	Gaps:	17

US-10-627-132-30 (1-1052) x AAZ99512 (1-3746)

Qy 34 GlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPhe 53

Db 366 GGCAGGTGTGCCAGATCTCGGCGACGGCGTGGCCACACGGGAGGGGCGGCTTC 425
Qy 54 ValAlaCysAenGluCysGlyPheProValCysArgProCysTyrGluTyrGluArg 73
Db 426 GCGCCTGCGACGCTCGCGGTTTCGGGTGTCGCGCCCTGCTACGAGTACGAGCGCAAG 485
Qy 74 GluGlyThrGlnAenCysProGlnCysGlyThrArgTyrCysArgLeuLysGlySerPro 93
Db 486 GACGGCACGGCGGTGCGCCCGACGTGCAAGACCAAGTACAAAGCGCCCAAAAGGGAGCGCG 545
Qy 94 ArgValAlaGlyAenAaspAenGluGluAaspAenGluHleGluPheAen--- 112
Db 546 GCGATCCGTGGG-----GAGGAGGAGACGACACTGATGCCGATACGACTTCAATTAC 599
Qy 113 IleAaspAenGluAenGlnArgGlnLeuGluGlyAenMetGlnAenSerGlnIleThr 132
Db 600 CTTGCATCTGGCAATGAGGAC-----CAGAAGCAGAGATTGCC 638
Qy 133 GluAlaMetLeuHleGlyArgMetSerTyrGlyArgGlyProAaspAenGly----- 149
Db 639 GACAGAAATGCGACGCTGCGCGCATGAACGTTGGGGGCGACGGGGATGTTGGTGGCCCCAAG 698
Qy 150 -----AaspGlyAenAenThrProGln---Ile 157
Db 699 TATGACAGTGGGAGATCGGGCTTACCAAGTATGACAGTGGCGAGATTCTCGGGGATAC 758
Qy 158 ProProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAen 177
Db 759 ATCCCATCAGTCACTAAACAGCCAG-----ATCTCAGGAGAAATCCCTGGTGTCTCC 809
Qy 178 GlyTyrGlyHis---GlyGluValSerSerSerLeuHleHisAglle---HisProTyr 195
Db 810 CCTGACCATCATATGATGTCCCCAACCTGGGAACATTGGCAAGCGTGTCCATTTCCTAT 869
Qy 196 ProValSerGluProGlySerAlaLysTrpAaspGluLysLys-----GluValSerTrp 213
Db 870 GTGAACCATTCGCCAAATCCGTCAGGAGGATCTCTGTTAGCATTTGGGAATGTTGCTGG 929
Qy 214 LysGluArgMetAaspTrpLysSerLysGln-----GlyIleLeu----- 227
Db 930 AAAGAGAGGGTGTATGGTGTGAAAAATGAAGCAGGACCAAGGGGAGCATTCCTCAGCAAT 989
Qy 228 -----GlyGlyGlyAlaAaspProGluAaspMetAenAla----- 238
Db 990 GGCACAGCATGTCTCCCTCTGAGGGTCGGGGTGTGGTGATATTGATGATCAACTGAT 1049
Qy 239 -----AaspValAlaLeuAenAaspGluAlaArgGlnProLeuSerArgLysVal 254
Db 1050 TACAACATGGGAAGATCGCTTATTGACCGCGAAACTCGACAGCCTCTATCTAGGAAGTT 1109
Qy 255 SerIleAlaSerSerLysValAenProTyrArgMetValIleValArgLeuValVal 274
Db 1110 CCACCTCTCTCTCCAGGATAAATCCATACAGGATGGTCAITGTGCTCGGATTGATTGT 1169
Qy 275 LeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAaspAlaIleGlyLeuTrp 294
Db 1170 CTAAGCATCTCTTGTGACATACCGTATCAAAATCCTGTGGCAATGACATCCCATATTGG 1229
Qy 295 LeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAaspGlnPhePro 314
Db 1230 CTTCTATCTGTATATGTGAGATCTGGTTGCTCTTTCGTGGATATTGGATCAGTTCCT 1289
Qy 315 LysTrpPheProIleAaspArgGluThrTyrIleAaspArgLeuSerLeuArgTyrGluArg 334
Db 1290 AAGTGGTTTCCAATCAACGGGAGACGTACCTTGATAGGCTGGCATTAAAGGTATGACCGG 1349
Qy 335 GluGlyGluProSerLeuSerAlaValAaspLeuPheValSerThrValAaspProLeu 354
Db 1350 GAAGGTGAGCCATCTCAGTTGGCTGCTGTGTACATTTTCGTGAGTACAGTGCACCAATG 1409
Qy 355 LysGluProProLeuValThrAlaAenThrValLeuSerIleLeuAlaValAaspTyrPro 374

Db 1410 AAGGAGCCTCTCTTGTGCTACATACCGTGCTATCCATTCTTGTGTGGATTACCCT 1469
Qy 375 ValAaspLysValSerCysTyrValSerAenAaspGlyAlaSerMetLeuThrPheGluSer 394
Db 1470 GTGGATAAGGTCTCTTGTCTATGATCTGATGATGGAGCTCGGATGCTGACATTTGATGCA 1529
Qy 395 LeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIle 414
Db 1530 CTAGCTGAGACTTCAGAGTTTGTAGAAAATGGGTACCATTTTGTAAAGATGACAAACATT 1589
Qy 415 GluProAraGalaProGluPheTyrPheSerLeuLysValAaspTyrIleLysAaspLysVal 434
Db 1590 GAACCTAGAGCTCCTCAATGGTACTTCTCCACAGAAAATGATTTACTTGAAGGACAAAGTG 1649
Qy 435 GlnProThrPheValGlnArgArgAlaMetLysArgGluTyrGluGluPheLysVal 454
Db 1650 CACCTCTCAATTTTAAAGACCGCGCGGCATGAGAGAGAATATGAAGAATTTCAAAGTT 1709
Qy 455 ArgIleAenAlaLeuValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLys 474
Db 1710 AGGTTAATGGCCTTGTGTAAGGCACAGAAAATTCCTGAGGAGGATGATCATGCAA 1769
Qy 475 AaspGlyThrProTrpProGlyAenAenThrArgAaspHisProGlyMetIleGlnValPhe 494
Db 1770 GATGGCACACCATGGCGAGGAAACAATACCMGGGACCATCTCGGAATGATTCAGGTTTTTC 1829
Qy 495 LeuGlyHisSerGlyGlyHisAaspThrGluGlyAenGluLeuProArgLeuValTyrVal 514
Db 1830 CTTGGTGCACAGTGGTGGCTTGTGTAAGGCACAGAAAATTCCTGAGGAGGATGATCATGTT 1889
Qy 515 SerArgGluLysArgProGlyPheGlnHisLysLysAlaGlyAlaMetAenAlaLeu 534
Db 1890 TCTCGTGAAGAAGGCTCCTGGATTCAGCATCACAGAAAGCTGGTGCATGAATGCTCTT 1949
Qy 535 IleArgValSerAlaValLeuThrAenAlaProPheMetLeuAenLeuAaspCysAaspHis 554
Db 1950 GTTCTGTCTCAGCTGTGCTTACCAATGGCAATACATGTTGTAATCTTGTGATGATCAC 2009
Qy 555 TyrIleAenAenSerLysAlaIleArgGluAlaMetCysPheLeuMetAaspProGlnVal 574
Db 2010 TACATTAAACAACAGTAGGCTCTCAGGGAAGCTATGTGCTTCTTATGGACCTTAACCTA 2069
Qy 575 GlyArgLysValCysTyrValGlnPheProGlnArgPheAaspGlyIleAaspValHisAasp 594
Db 2070 GGAAGGAGTCTCTGCTACGTCAGTTTCCCGACAGATTCGATGGCATTTGACAGGAATGAT 2129
Qy 595 ArgTyrAlaAenAraAenThrValPhePheAaspIleAenMetLysGlyLeuAaspGlyIle 614
Db 2130 CGATATGCCAACAGGAACACCGTGTGTTTTCGATATTAACTTGAGAGGTCTTGTATGGCATC 2189
Qy 615 GlnGlyProValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyr 634
Db 2190 CAAGGACAGTTTATGCGGAACCTGGCTGTGTTTTCACCGAACAGCTCTATATGTTAT 2249
Qy 635 AenProProLysGlyProLysArgProLysMetValThrCysAaspCysCysProCysPhe 654
Db 2250 GAGCCCCCAATTAGCAGAGAGAGGGTGGTTCTTGTGCA-----TCACTATGTGGC 2300
Qy 655 GlyArgLysLys-----ArgLysHis 661
Db 2301 GGTAGGAAGAAGGCAAGCAAAATCAAAGGGCTCGGACAGAAAGATGCCAGAAAGCAT 2360
Qy 662 AlalysAaspGlyLeuPro-----GluGlyThrAlaAaspMet 673
Db 2361 GTGACAGATCTGTGGCCAGGTATTCACTTGAAGATATAGAGGAGGAGTTGAAGGCGCT 2420
Qy 674 GlyValAaspSerAaspLysGluMetLeuMetSerHisMetAenPheGluLysAraPheGly 693
Db 2421 GGATTTGACGACGAGAAATCACTTCTTATGCTCAAATGAGCCTGGAGAGAGATTGGC 2480
Qy 694 GlnSerAlaAlaPheValThrSerThrIleuMetGluGlyGlyValProProSerSer 713
Db 2481 CAGTCCGAGCGGTTTGTTCCTCCACTCTGATGGAGTATGGTGTGTCTTCCTCAGTCCGCA 2540

```
QY 714 SerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLys 733
Db ACTCCGAGTCTCTTCTGAAGAAGCTATCCATGTTATAGCTGTGGCTATGAGACAAG 2600
QY 734 ThrAspTrpGlyLeuGlyLeuGlyTyrPileTyrGlySerIleThrGluAspIleLeuThr 753
Db ACTGAATGGGAACTAGATCGGTGATCTACCGGTCTGTGACAGAAGACATTCACCC 2660
QY 754 GlyPheLysMetHisCysArgGlyTyrArgSerValTyrCysMetProLysArgAlaAla 773
Db CGATTCAAGATGCACGCGAGGCTGGCGTGCATCTACTGATGCCCCAAGCGCCAGCT 2720
QY 774 PheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAla 793
Db TTCAAGGGGTCTGCCCCCATCAATCTTCGGACCGCTGAACACGAGTGTCCGGTGGCT 2780
QY 794 LeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsn 813
Db CTTGGTCTCGTGAGATCTCTTCACCGCGCATCTGCCCTGTGGTACGGCTAC---GGA 2837
QY 814 GlyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThr 833
Db GGGCGGCTCAAGTCTCTGGAGATTCGGGTATCATCAACACCATCTACCGCTCAGC 2897
QY 834 SerLeuProLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPhe 853
Db TCCATCCCGCTTCTCATCTACTGCTCTGCCCGCATCTGTCTGTCTACCGGAAAGTTC 2957
QY 854 IleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIle 873
Db ATCATCCAGATCAGCAACTTCGCCAGCATCTGGTTTCATCTCCCTCTTCATCTCGATC 3017
QY 874 PheAlaThrGlyLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArg 893
Db TTCCGCCACGGCATCTCGAGATGAGTGGAGGGTGGGCATCGACAGTGTGGGAGG 3077
QY 894 AsnGluGlnPheTrpValIleGlyValSerAlaHisLeuPheAlaValValGlnGly 913
Db AACGAGCAGTCTTGGGTGATCGGGGCATCTCCGGCAGCTCTTCGCGTGTTCAGGGC 3137
QY 914 LeuLeuValValLeuAlaGlyIleAspThrAsnPheThrValThrSerIleAlaThrGly 933
Db CTGCTCAAGGTGTGCCCGCATCGACACCACTTCACCGTCACTCCCAAGGCTCG--- 3194
QY 934 AspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrLeuLeuIlePro 953
Db GACGAGGACGGCACTTCGGGAGCTGTACATGTTCAAGTGACGAGCGTCTCTGATCCG 3254
QY 954 ProThrThrLeuLeuIleAsnValIleGlyValValAlaGlyIleSerAspAlaIle 973
Db CCCACCCACCATCTCATCATCAACTGTCGGCGTCTCGCGGCATCTCTTACGCCATC 3314
QY 974 AsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaThrTrpVal 993
Db AACACGGGATACAGTCGGGGCGCGCTTCGCGCAAGCTCTTCTGCGCTTCGGGTC 3374
QY 994 IleValHisLeuTyrProPheLeuLysGlyLeuMetGlyValArgGlnAsnArgThrProThr 1013
Db ATCGTCCACCTGTACCGCTTCTCAAGGCGCTCATGGGCGAGCAGACCGGACCCCGACC 3434
QY 1014 ValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIle 1033
Db ATCGTGTCTGTGGGCCATCTCTGTGGGTCTCATCTTCTCTGTGTGGGTTCGCATC 3494
QY 1034 AspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db GACCCCTTCCACCACCGCGTCACTGCGCGCGATACCCAGACGTGTGGCATCAACTGC 3551
RESULT 15
AAZ99527
ID AAZ99527 standard; DNA; 3746 BP.
XX
```

```
AC AAZ99527;
XX 03-JUL-2000 (first entry)
XX DNA encoding a maize cellulose synthase.
XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
XX transgenic plant; plant breeding marker; ss.
XX Zea mays.
XX Key Location/Qualifiers
XX CDS 321..3449
XX /product= "cellulose synthase"
XX /tag= a
XX /transl_except= (pos: 1800..1802, aa: Xaa)
XX /note= "no termination codon given; Xaa is an unspecified
XX amino acid"
XX WO200009706-A2.
XX 24-FEB-2000.
XX 16-AUG-1999; 99WO-US018760.
XX 17-AUG-1998; 98US-0096822P.
XX (PION-) PIONEER HI-BRED INT INC.
XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX WPI; 2000-224343/19.
XX P-PSDB; AAY84119.
XX New genes which encode maize cellulose synthase polypeptides in plants
XX useful for modulating the expression of cellulose synthase in plants and
XX to produce transgenic plants expressing the novel protein.
XX Claim 1; Page 176-181; 119pp; English.
XX The present sequence encodes a maize cellulose synthase polypeptide. The
XX cellulose synthase can be used for the improvement of stalk quality for
XX improved stand or silage. It also provides an increased concentration of
XX cellulose in the pericarp, hardening the kernel and improving its
XX handling ability. The sequences are used to produce transgenic plants and
XX seeds expressing the cellulose synthase. The polynucleotide is used for
XX modulating, preferably increasing, the level of the synthase in a plant
XX cell. The plants are preferably monocots. The polynucleotide is also used
XX as a probe or primer in the detection quantitation or isolation of gene
XX transcripts. The probes are useful in detecting deficiencies in the level
XX of mRNA in screenings for desired transgenic plant, for detecting or
XX mutations in the gene, for monitoring upregulation of expression or
XX changes in enzyme activity in screening assays of compounds, for
XX detection of any number of allelic variants of the gene, or for use as
XX molecular markers in plant breeding programs. The isolated nucleic acids
XX of the present invention can also be used for recombinant expression of
XX their encoded polypeptides or for use as immunogens in the preparation
XX and/or screening of antibodies. The proteins can be employed in assays
XX for enzyme agonists or antagonists of enzyme function or for use of
XX immunogens or antigens to obtain antibodies specifically immunoreactive
XX with a protein
XX SQ Sequence 3746 BP; 861 A; 976 C; 1006 G; 902 T; 0 U; 1 Other;
Alignment Scores:
Pred. No.: 0 Length: 3746
Score: 3912.50 Matches: 750
Percent Similarity: 78.96% Conservative: 102
Best Local Similarity: 69.51% Mismatches: 150
Query Match: 69.27% Indels: 77
DB: 3 Gaps: 17
US-10-627-132-30 (1-1052) x AAZ99527 (1-3746)
```

QY 34 GlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPhe 53
Db 366 GGACAGGTGTCAGATCTCGCGGACCGCGTGGCCACCGCGGAGGGGAGCTCTTC 425
QY 54 ValAlaCysAenGluCysGlyPheProValCysArgProCysTyrGluTyrGluArg 73
Db 426 CCCGCTCGACGCTCGGGGTTCCGGTGTCCGCCCTCTACGAGTACGAGCGCAAG 485
QY 74 GluGlyThrGlnAenCysProGlnCysLysThrArgTyrLysArgLeuLysGlySerPro 93
Db 486 GACGCGACGCGAGCGTGCCTCCAGTGCAGCAAGTACAAAGCGCCCAAGGGAGCGCG 545
QY 94 ArgValAlaGlyAspAspGluClnAspIleAspLeuGluHisGluPheAen--- 112
Db 546 GCGATCCCTGGG-----GAGAGGAGGACGACACTGATGCCGATGACCTTCAATTAC 599
QY 113 IleAspAspGluAenGlnArgGlnLeuGluGlyAenMetGlnAenSerGlnIleThr 132
Db 600 CTTCATCTGGCAATGAGGAC-----CAGAAGCAGAAGATTGCC 638
QY 133 GluAlaMetLeuHisGlyArgMetSerTyrGlyArgGlyProAspAspGly----- 149
Db 639 GACAGAATGCGCAGCTGCGCGATCAACGTTGGGGGACGCGGGGATGTTGGTCCGCCCAAG 698
QY 150 -----AspGlyAenAenThrProGln-----Ile 157
Db 699 TATGACATGCGGAGATCGCGCTTACCAAGTATGACAGTGGCGAGATTCTTCGGGGATAC 758
QY 158 ProProIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAen 177
Db 759 ATCCCATCAGTCACTAACAGCGAG-----ATCTCAGGAGAAATCCCTGGTGTCTCC 809
QY 178 GlyTyrGlyHis---GlyGluValSerSerSerLeuHisLysArgIle---HisProTyr 195
Db 810 CCGTACCATCATATGATGATCCCAACTGGGAACATTGGCAAGCGGTCTCCATTTCCTAT 869
QY 196 ProValSerGluProGlySerAlaLysTrpAspGluLysLys-----GluValSerTrp 213
Db 870 GTGAACCAATGCGCAAAATCCGTCAGGAGATTCTCGTAGCATTTGGGAATGTTGCCCTGG 929
QY 214 LysGluArgMetAspAspTrpLysSerLysGln-----GlyIleLeu----- 227
Db 930 AAAGAGAGGTTGATGGCTGGGAAATGAAGCAGGACAAAGGCGAGATTCCCATGACGAAT 989
QY 228 -----GlyGlyAlaAspProGluAspMetAspAla----- 238
Db 990 GGCACAAGCATTTGCTCCCTCTGAGGGTGGGGTGTGGTGATATTGATGTCATCAACTGAT 1049
QY 239 -----AspValAlaLeuAenAspGluAlaArgGlnProLeuSerArgLysVal 254
Db 1050 TACAACATGGAAGATGCGCTTATTGAACGACGAACTCGACAGCCCTCTATCTAGGAAAGTT 1109
QY 255 SerIleAlaSerSerLysValAenProTyrArgMetValIleValValArgLeuValVal 274
Db 1110 CCATCTCTCTCCAGGATAAATCCATACAGAGTGGTCATTGTCTGCGATTGATTGTT 1169
QY 275 LeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrp 294
Db 1170 CTAAGCATCTCTTGCACTACCGTATCAAAATCTGTGCGCAATGTCATCCCATTTATGG 1229
QY 295 LeuValSerIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPhePro 314
Db 1230 CTTCATCTGTTATATGAGATCTGTTGCTCTTTCTGCGATATTGATCAGTTCCCT 1289
QY 315 LysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArg 334
Db 1290 AAGTGGTTTCCAATCAACGGGAGAGCTACCTTTGATAGCTGGCATTAAGGTATGACCGG 1349
QY 335 GluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSerThrValAspProLeu 354
Db 1350 GAAGGTGAGCCATCTCAGTTGGCTGCTGTGATCTTTCTGTCAGTACAGTCGACCCAAATG 1409

QY 355 LysGluProProLeuValThrAlaAenThrValLeuSerIleLeuAlaValAspTyrPro 374
Db 1410 AAGAGCCTCTCTTGTCACTGCAATACCGTGTATCCATTCTTGTGTGGATTACCTT 1469
QY 375 ValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheGluSer 394
Db 1470 GTGGATAAGGCTCTTGTCTATGATCTGATGATGAGCTCGGATGTCGATTTGATGCA 1529
QY 395 LeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIle 414
Db 1530 CTAGCTGAGACTTCCAGAGTTTGTAGAAATGGGTACCATTTGTTAAGAAGTACAACTT 1589
QY 415 GluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysVal 434
Db 1590 GAACCTAGACTCTCTGAATGGTACTTCTCCAGAAAATTGATTACTTTGAAGGACAAAGTG 1649
QY 435 GlnProThrPheValGlnGluArgAlaMetLysArgGluTyrGluGluPheLysVal 454
Db 1650 CACCTTCAATTTGTTAAAGACCCCGCGGCCCATGAAGAGAAATATGAAGAATTCAAAGTT 1709
QY 455 ArgIleAenAlaLeuValAlaLysAlaMetLysValProAlaGluGlyTyrIleMetLys 474
Db 1710 AGGGTAATAGCCCTTGTGTTAGGCGACAGAAAGTTCTCAGGAGAGATGGATCATGCAA 1769
QY 475 AspGlyThrProTrpProGlyAenAenThrArgAspHisProGlyMetIleGlnValPhe 494
Db 1770 GATGGCACACATGGCCAGGAAACAATACCMGGGACCATCTCGAATGATTCAGGTTTC 1829
QY 495 LeuGlyHisSerGlyGlyHisAspThrGluGlyAenGluLeuProArgLeuValTyrVal 514
Db 1830 CTTGGTCACAGTGGTGGCTTGATCTAGGCGCAATGAGCTACCCCGCTTGGTCTATGTT 1889
QY 515 SerArgGluLysArgProGlyPheGlnHisLysLysValAlaGlyAlaMetAsnAlaLeu 534
Db 1890 TCTCGTGAAGAGCGTCTCGGATCCAGCATCACAAAGAGCTGGTGCATGAATGCTCTT 1949
QY 535 IleArgValSerAlaValLeuThrAenAlaProPheMetLeuAenLeuAspCysAspHis 554
Db 1950 GTTCGTCTCTCAGCTGTGCTTACCAATGGACAATACATGTTGAATCTTGTGATGATCAC 2009
QY 555 TyrIleAenAenSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnVal 574
Db 2010 TACATTAAACACAGTAAGGCTCTCAGGGAAGCTATGCTTCTTATGGACCCCTAAACCTA 2069
QY 575 GlyArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAsp 594
Db 2070 GGAAGGAGTGTCTGCTACGTCAGCTTCCCGAGAGATTCGATGCGCATTCAGAGAAATGAT 2129
QY 595 ArgTyrAlaAenArgAenThrValPhePheAspIleAenMetLysGlyLeuAspGlyIle 614
Db 2130 CGATATGCCAACAGAACACCGCTGTTTTTCGATATTAACTTGAGAGGTCTTGTATGGCATC 2189
QY 615 GlnGlyProValTyrValGlyThrGlyCysValPheArgGlnAlaLeuTyrGlyTyr 634
Db 2190 CAAGGACAGTTTATGTCGGAACCTGGTGTGTTTCAACCGAACAGCTCTATATGGTTAT 2249
QY 635 AsnProProLysGlyProLysArgProLysMetValThrCysAspCysAspProCysPhe 654
Db 2250 GAGCCCCCAATTAAGCAGAGAGGGGTGTTCTTGTGCA-----TCATATGTGGC 2300
QY 655 GlyArgLysLys-----ArgLysHis 661
Db 2301 GGTAGGAAGAGGCAAGCAAAATCAAGAGGGCTCGGACAAGAAAGTCCGAGAAGCAT 2360
QY 662 AlaLysAspGlyLeuPro-----GluGlyThrAlaAspMet 673
Db 2361 GTGGACAGTCTCTGCGCAGATATTCAACCTTGAAGATATAGAGGAGGGAGTTGAAGCCCT 2420
QY 674 GlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGly 693
Db 2421 CGATTGACGAGAGAAATCACTTCTTATGTCTCAATGACCTCGGAGAGAGATTGGC 2480
QY 694 GlnSerAlaAlaPheValThrSerThrLeuMetGluGlyGlyValProProSerSer 713

This Page Blank (uspto)

504 CAG-----TCTGTGGCCGAGTCCATGCTCTACGGC 533
139 ArgMetSerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIlePro 158
534 CACATGAGCTACGCCCGTGA-----GGTGACCTTAATGGCGGCCACAGCTTTC 584
159 ProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGly 178
585 CAGTC-----AACCCCAATGTTCCACTC-----CTCACCAACGGG 620
179 TyrGlyHisGlyGluValSerSerLeuHis----- 189
621 CAAATGTGGATGACATCCACCGAGCAGCAGCGCTGTGGCTTCTTTCATGGGTGGT 680
190 -----LysArgIleHisProTyrProValSerGluProGly-----SerAla 203
681 GGGGAAAGAGGATACATCCCTTCTTATGCGGATCCAGCTTACCTGTGCAACCCAGG 740
204 LysTrpAspGluLysGlu-----ValSerTrpLysGlu 215
741 TCTATGACCCATCCAAAGGATCTGTGTCATATGGGTATGGTAGTGTTGCTTGGAGGAA 800
216 ArgMetAspAspTrpLysSerLysGln-----GlyIleLeuGlyGly 229
801 CGGATGAGGAATTCGAAGCAGAGCAAGAGAGGATGCACAGACGGCGGAATGATGGTGGT 860
230 GlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnPro 249
861 GGTGATGATGGTGAC-----GATGCTGATCTACCACTAATGGATGAAGCAAGCAACA 914
250 LeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleVal 269
915 CTGTCCAGGAATAATCCATTCATCAAGCAGCAATTAATCCATATAGATGATATCATY 974
270 ValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAsp 289
975 ATTCGGCTTGTGGTGTCTTCTTCCACTACCGAGTGATCGCGTGAATGAT 1034
290 AlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIle 309
1035 GCATTTGCTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTGGCATGTCTTGGATT 1094
310 LeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSer 329
1095 CTTGATCAATTCACCAAGTGTTCCTATTGAGAGAGACTTACCTAGACCGCTGTCA 1154
330 LeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSer 349
1155 CTGAGTTTCGACAAGGAGCCGCACTCTCAACTTGTCTCCAAATGATTTCTTGTGAGT 1214
350 ThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeu 369
1215 ACGTGTGATCCCTTAAAGGAACCTCTTGTGTCACAAATACTGTTCTATCTATCCTT 1274
370 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 389
1275 TCGTGGATTAATCCCTGTTGATAGGTTTCTGCTATGTTTCTGATGATGGTGTGCAATG 1334
390 LeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCys 409
1335 CTAACGTTTGAAGCATATATCTGAAACATCTGAAATTTGCAAGAAATGGTTCTTCTTCG 1394
410 LysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyr 429
1395 AAACCGGTACAAATATTGAACCTCGCGCTCCAGGTGGTACTTCCAAACAGAGATAGACTAC 1454
430 LeuLysAspLysValGlnProThrPheValGlnArgArgAlaMetLysArgGluTyr 449
1455 TTGAAGACACAGGTGGCAGCAACTTGTAGGGAGAGAGAGCAATGAAGAGAGATAT 1514
450 GluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGlu 469
1515 GAGGAATTCAGGTGAGAAATCAATGCCTTAGTTGTCAAAGGCCCAAGAAAGTTCTCTGAAGA 1574

470 GlyTrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly 489
1575 GGNATGACAAATGCAAGATGGAAACCCCTGGCCTGGGAAACAATGTTCTGTATCATCTCTGA 1634
490 MetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuPro 509
1635 ATGATTCAAGTCTTCTCTGGCCAAAGCGGAGGCTTGTGCTGTGAGGAAATGAAGTCCCA 1694
510 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly 529
1695 CGATTGGTTATGTTTCTAGAGAGAAACGACCAAGGCTATAACCATCATAGAAGAGCTGT 1754
530 AlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsn 549
1755 GCTATGAATGCAATGGTCCGAGTCTCTGCTGTACTAAACAATGCTCCATATTTGTTAAAC 1814
550 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeu 569
1815 TTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTATG 1874
570 MetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGly 589
1875 ATGGACCTTACTAGGAAAGAGGTTTGTATGTACAGTCTCCTCAAGATTTGATGGG 1934
590 IleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLys 609
1935 ATTGATCGCATGACCGATATGTAACCGGAATGTTGTCTTTTGTATATCAACATGAA 1994
610 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGln 629
1995 GGTTCGATGGTATTTCAGGGTCCAAATTTATGTTGGTACTCGATCTGTATTTAGAACGAG 2054
630 AlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysAsp 649
2055 GCATTATATGTTATGATGCCCCCAAA---ACAAGAAAGCCACCATCAAGAGCTTGCAAC 2111
650 -----CysCysProCysPheGlyArgLysLysArgLysHisAla 662
2112 TGCTGGCCCAAGTGGTGTCTTTGCTGTGCTGCTTTGGCAATAGGAAGCAAGAGAGACT 2171
663 LysAspGlyLeuProGlu-----GlyThrAlaAspMetGlyValAspSerAspLysGlu 680
2172 ACCAAACCCCAAAACAGAGAAAGAAAGTTATTTTCAAGAAAGAGAAACCAATCC 2231
669 -----GlyThrAlaAspMetGlyValAspSerAspLysGlu 680
2232 CCTGCATATGCTCTTGTGTAATTTGCAAGAGCTGCTCCAGAGCTGAGAATGAAGAGCC 2291
681 MetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThr 700
2292 GGTATGTTAATCAACAAAATTAGAAAGAAATTTGCCCAATCTCTGTTTTTTGTATCA 2351
701 SerThrLeuMetGluGluGlyValProProSerSerSerProAlaAlaLeuLeuLys 720
2352 TCCACACTTCTCGAGAATGGTGAACCTTCAAGAGTGCAGTCTCTCTCTTTTGA 2411
721 GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeu 740
2412 GAAGCTATACATGTCATGTGTTGTTGTTATGAAGACAGACAGACTGGGGAAGAGAT 2471
741 GlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg 760
2472 GGCTGGATCTATGATCAGTTACAGAGATATTTCTAATCGTTTTCAAGATGCAATGTCTAT 2531
761 GlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIle 780
2532 GGTGGCGGTCAATTTACTGCATACCTAAACGGGTTCATTTCAAAAGGTTCTGCACCTCTG 2591
781 AsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePhe 800
2592 AATCTTTCAGATCTCTTCCACCGTCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTC 2651

```
QY 801 PheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGlu 820
Db 2652 TTCAGCAATCATGGCCCTCTTTGGTATGGGTAT---GGTGGCGGTCTGAAATTTTGGAA 2708
QY 821 ArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuLeuLeuAlaTyr 840
Db 2709 AGATTTCTCATCATCACTCCATCCATCGTGTATCTCTGGACATCTATTCCTCTCTGGCTTAC 2768
QY 841 CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr 860
Db 2769 TGTACATTGCCTGCCATCTGTTTATTTACAGCGAAATTTATCACTCCAGAGCTGAATAAT 2828
QY 861 PheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGlu 880
Db 2829 GTTGCAGCCTGTGGTGTATGTCATCTTTTATCTGATTTTGTCTACGAGCATCTCAGAA 2888
QY 881 MetArgTrpSerGlyValSerIleGluGluTyrTrpArgAsnGluGluPheTrpValIle 900
Db 2889 ATGAGATGGAGTGGTGTGGATTCGATGACTGTGTGGAGGATGACAGATTCTGGGTCAAT 2948
QY 901 GlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAlaGly 920
Db 2949 GGAGGTGTCTCTCACACCTCTTGTCTGTGTCCAGGACCTTCTCAAGTCTACACTGCTGT 3008
QY 921 IleAspThrAsnPheThrValThrSerIysAlaThrGlyAspGluAspGluPheAla 940
Db 3009 GTTGATACAAGTTCACCGTGCATCAAAAG-----GGTGGAGATGATGAGGAGTCTCA 3062
QY 941 GluLeuTyrAlaPheLeuTyrTrpThrLeuLeuIleProProThrThrLeuLeuIleIle 960
Db 3063 GAGCTATATACATTCAAATGGACTACCTTATGATACCTCTTACCACCTTGTCTTATGT 3122
QY 961 AsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTrp 980
Db 3123 AACTTCATTGGTGTGTGCTGCTGGCGTTTCAATGCGATCAATAACGATATGAGTCATGG 3182
QY 981 GlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPhe 1000
Db 3183 GGCCCCCTCTTTGGGAAGCTATCTTTCGATTTTGGGTGGATTGTCATCTTTATCCCTTT 3242
QY 1001 LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValValIleTrpSerIle 1020
Db 3243 CTCAAAGTGTGGTGGAGGCAAAACAGGACCAACGATTCATGCTCTGTGTCAT 3302
QY 1021 LeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThr 1040
Db 3303 CTGCTGGCTTCAATCTCTCGCTCTTGTGGTGGATTGATCTTCTTCTTCCGGAAGAT 3362
QY 1041 LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3363 GATGTCCTCGCTCTTGGAGGAGTGTGGTTGGATTGC 3398
```

RESULT 3

```
US-10-160-719A-1
; Sequence 1, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugaa, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
```

```
; LENGTH: 3568
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (63)...(3239)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3487)...(3487)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(3568)
; OTHER INFORMATION: n = A,T,C or G
US-10-160-719A-1
```

Alignment Scores:

Pred. No.:	0	Length:	3568
Score:	4005.00	Matches:	743
Percent Similarity:	79.04%	Conservative:	128
Best Local Similarity:	67.42%	Mismatches:	141
Query Match:	70.91%	Indels:	90
DB:	4	Gaps:	16

US-10-627-132-30 (1-1052) x US-10-160-719A-1 (1-3568)

```
QY 14 ArgAsnGluLeuValLeuIleArg-----GlyHisGluAspProLysProLeuArgAla 31
Db 12 CGTCCGAGCTCGTCTGTCATCCGCGCATGCGAGCCAGGCGCGAAGCCCATGAGCCAG 71
QY 32 LeuSerGlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAsp 51
Db 72 CGGAACGCCAGGTGTGCCAGATTTCGGCGACGACGCTGGGGCGCAACCCCGACGGGAG 131
QY 52 LeuPheValAlaCysAsnGluCysGlyPheProValCysArgProCysTyrGluTyrGlu 71
Db 132 CTTTGTGGCTCGAACAGTGTGCGCTTCCCATCTGCGGGACTGTCTAGTACGATCCAG 191
QY 72 ArgArgGluGlyThrGlnAsnCysProGlnCysLysThrArgTyrLysArgLeuLysGly 91
Db 192 CGCCCGAGGCGACGCGAAGTGCCTCCAGTGCAGACCGCTTCAAGCGCTTCAAGGGG 251
QY 92 SerProArgValAlaGlyAspAspGluAspLeuAspIleAspAspLeuGluHisGluPhe 111
Db 252 TGCCTGCGCGTGGCGGGGACGAGGAGGAGGAGCGCTCGACGACCTTGGAGAACGAGTTC 311
QY 112 AsnIleAspAspGluAsnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIle 131
Db 312 AACTGGAGCGACAGCAGTCCAGTACCTC----- 344
QY 132 ThrGluAlaMetLeuHisGlyArgMetSerTyrGlyArgGlyProAsp---AspGlyAsp 150
Db 345 GCGAGTCCATGCTCCACGCCACATGAGCTACGCGCGCGCGCGGACCTCGACGGCGTG 404
QY 151 GlyAsnAsnThrProGlnIleProProIleIleThrGlySerArgSerValProValSer 170
Db 405 CCGAGCGCATTCACCCCATCCCTC-----AATGTTCCCTC--- 440
QY 171 GlyGluPheProIleThrAsnGlyTyrGlyHisGlyGluValSerSerSerLeuHis--- 189
Db 441 -----CTCACCAACGACAGATGTCATGTCATCCCGCGGAGCAGCACGCC 488
QY 190 -----LysArgIleHisProTyrProValSerGlu 199
Db 489 CTTGTGCCCTCGTTCTGGTGGTGGGGGGAAGAGGATTACACCTCTCCCTACCGCAT 548
QY 200 ProGly-----SerAlaLysTrpAspGluLysLysGlu----- 210
Db 549 CCCAACCTTCCTGTGTCAACCGAGGTCTATGACCCCTTCCAGGATCTCGCGCATATGCC 608
QY 211 -----ValSerTrpLysGluArgMetAspAspTrpLysSerLysGlnGlyIleLeu 227
Db 609 TACGGGAGCGTAGCATGGAAGGAGGATGAGAGTGGAGAGTGGGAAGCAGAGGAGGATG 668
```

Qy 228 -----GlyGlyGlyAlaAspProGluAspMetAspAlaAspValAla 241
Db 669 CACCAGACGAGAACGATGGCGGCGGATGGTGTAT-----GATGCAGATCTACCA 722
Qy 242 LeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSerSerLysVal 261
Db 723 CTAATGGATGAAGCTAGACAGCCATTGTCCAGAAAGATCCCGCTTCCTTCAAGCCAAATC 782
Qy 262 AsnProTyrArgMetValIleValArgLeuValValLeuAlaPhePheLeuAspTyr 281
Db 783 AACCCCTATAGGATGATTAATAATTCGGCTAGTGGTTTGTGTTCTTCTCCACTAC 842
Qy 282 ArgIleLeuHisProValProAspAlaIleGlyLeuTyrLeuValSerIleIleCysGlu 301
Db 843 CGAGTGTGATCCCGTGGCTGATGCTTTATGGCTCATATCTGTGATCTGTGAA 902
Qy 302 IleTyrPheAlaIleSerTyrIleLeuAspGlnPheProLysTyrPheProIleAspArg 321
Db 903 ATTTGGTTTGGCATGTCTTGGATCTTGACCAAGTTTCCAAAGTGGTTTCTTATCGAGAG 962
Qy 322 GluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGluProSerLeuLeu 341
Db 963 GAACCTATCTGTGACCGGCTGAGTTAAGGTTTGACAAAGAAAGGCGATCTTCTCAACTC 1022
Qy 342 SerAlaValAspLeuPheValSerThrValAspProLeuLysGluProProLeuValThr 361
Db 1023 GCGCTGTGTGATTTCTTGTGTCAGTACGGTTGATCCCTTGAAGAAACCTCCATTGGTCACT 1082
Qy 362 AlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyr 381
Db 1083 GCTAATAGTCTTATCTATCTCTTTCGGTGGATATCCAGTTGATGAAGGTTTCTATGCTAC 1142
Qy 382 ValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGluPhe 401
Db 1143 GTTTCTGATGATGGTGGCTGACATTTGAAGCATTTGTGAAACATCTGAAATTT 1202
Qy 402 AlaArgLysTyrValProPheCysLysLysPheGlyIleGluProArgAlaProGluPhe 421
Db 1203 GCNAAAGNATGGTTCCTTCTGCAAAAGATATAGCCTTGAGCCTCGTGTCTCCAGAGTGG 1262
Qy 422 TyrPheSerLeuLysValAspTyrIleLysAspLysValGlnProThrPheValGlnGlu 441
Db 1263 TACTTCCAAAGAGATAGACTACCTGAAAGACAAAGGTGGCGCCAAACTTTGTTAGAGAA 1322
Qy 442 ArgArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAlaLeuValAla 461
Db 1323 CGGAGCAATGAAGAGAGATGAGGAATTCAGAGTCAGAAATCAATGCCCTCTGTGTCT 1382
Qy 462 LysAlaMetLysValProAlaGluGlyTyrIleMetLysAspGlyThrProTyrProGly 481
Db 1383 AAAGCCCAAAAGTTCTCTGAGGAGGATGGACAAATGCAGGATGGAACCTCCATGCCCGGA 1442
Qy 482 AsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHis 501
Db 1443 AATAATGTCGGTATCATCCTGGAATGATTCAGGTTTTCCTTGGTCAAAAGTGGTGGCCAT 1502
Qy 502 AspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGly 521
Db 1503 GATGGGAAGGAAATAGAGCTGCTCGATGGTGGTTTATGTTTCAAGAGAAAAACGGCCAGGC 1562
Qy 522 PheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSerAlaValLeu 541
Db 1563 TACAACCATCACAGAGGCTGGTCTATGAATGATGATGGTCCGAGTCTCTGCTGACTA 1622
Qy 542 ThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAla 561
Db 1623 ACTAATGCTCTCTTATTTGTGAACCTTGGATTGTGATTCATATATCAATAATAGTAAGGCT 1682
Qy 562 IleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysValCysTyrVal 581
Db 1683 ATAAAGGAGCAATGTGTTTATGATGGATCTCTTGTGGNAGAAAGTTTGCTATG 1742
Qy 582 GlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsnArgAsnThr 601

Db 1743 CAGTTTCTCAAAAGATTTGATGGATTTGATCGCATATGATCGTATACAGAAATGTT 1802
Qy 602 ValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProValTyrValGly 621
Db 1803 GTCTTTTTCGATATCAACATGAAGAGTTTGGATGGTATCCAGGCCCAATTTATGTGGGT 1862
Qy 622 ThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProProLysGlyProLys 641
Db 1863 ACTGGATGTCTTCAGNAGCGCAGGCATATATATGGCTACGATCTCCCAA---ACAAAG 1919
Qy 642 ArgProLysMetValThrCysAsp-----CysCysProCysPhe 654
Db 1920 AAGCCACCATCAAGAACTTGCACACTGCTGGCCAAAGTGGTGCATTTGCTGTGTTT 1979
Qy 655 Gly---ArgLysLysArgLysHisAlaLysAspGlyLeuProGlu----- 668
Db 1980 GGTAAACAGGAAGAACCAAGAACAGCAACCTCTAAACCTAAATTTGAGAAGATAAAG 2039
Qy 669 -----GlyThr 670
Db 2040 AAACTTTTTAAAGAAAGGAAATCAAGCCCTGCATATGCTCTTGGTGAAATTTGATGAA 2099
Qy 671 AlaAspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLys 690
Db 2100 GCGCTCCAGAGCTGAAATGAAAGGCTAGTATTGTAATCAACACAGAAAGTTGGAAAG 2159
Qy 691 ArgPheGlnSerAlaPheValThrSerThrLeuMetGluGluGlyGlyValPro 710
Db 2160 AAATTTGGCCAGTCTTCAGTTTGTGTCATCCACACTTCTGAGAAATGGTGGNACCTG 2219
Qy 711 ProSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 730
Db 2220 AAGAGTGCCAGTCCAGCTTCTCTTGAAGAAAGCTATACATGTCATCAGTTGGATAT 2279
Qy 731 GluAspLysThrAspTyrGlyLeuGluLeuGlyTyrIleTyrGlySerIleThrGluAsp 750
Db 2280 GAAGACAAAACAGCTGGGGAAGATATTTGGTTGGATTTATGGATCAGTCACAGAAAGAT 2339
Qy 751 IleLeuThrGlyPheLysMetHisCysArgGlyTyrArgSerValTyrCysMetProLys 770
Db 2340 ATTTCTACTGGTTTAAAGATGCACTGCCATGGTTGGCGGTCAATTTACTGCATACCTAAA 2399
Qy 771 ArgAlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 790
Db 2400 CGGCGCCCTTCAAGGTTCCGACACTCTCAATCTTTCCGATCGTTTTCCACAGGTTCTT 2459
Qy 791 ArgTyrAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGly 810
Db 2460 CGGTGGGCTCTTGGTTCAATTTGAAATTTGTTTCAGCAACCACTGCCCTCTCTGATATGG 2519
Qy 811 TyrLysAsnGlyAsnLeuLysTyrLeuGluArgPheAlaTyrIleAsnThrThrIleTyr 830
Db 2520 TAT---GGTGGTGACATAAAGTTCTCGAAAGGTTTTCGTACATTAATCTCATCGTATAC 2576
Qy 831 ProPheThrSerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThr 850
Db 2577 CTTGGACATCTATCCCGCTCTTGGCTATTTGCACATTTGCTGCTCATCTGCTGACA 2636
Qy 851 GlyLysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPhe 870
Db 2637 GGGAAATTTATCAGCCAGCAGCTTAAACAATGTTGCCAGCCTCTCGTTTCATGCTCATCTTC 2696
Qy 871 MetSerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValSerIleGluGlu 890
Db 2697 ATCTGCATTTTGTCTAGCAGCATCTCGAATAGATGGAGTGGTGTAGGCATCGATGAC 2756
Qy 891 TrrTyrArgAsnGluGlnPheTyrValIleGlyValSerAlaHisLeuPheAlaVal 910
Db 2757 TGGTGGAGAAACGAGCAGTTTGGTCTATTGGAGCGGTGCTCTCACATCTCTTGTCTGTG 2816
Qy 911 ValGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLys 930

```
Db      2817  TTCCAGGGACTCTCAAGGTCTAGCTGGTGTAGACACGAGCTTCACTGTGACATCCAAG 2876
Qy      931  AlaThrGlyAspGluAspGluPheAlaGluLeuTyzAlaPheIysTrpThrThrLeu 950
Db      2877  -----GGCGGAGACGACGAGGAGTTCTCAGAGCTGTACACATTCAATGGACGACCTT 2930
Qy      951  LeuIleProProThrThrLeuLeuIleIleAsnValIleGlyValIleAlaGlyIleSer 970
Db      2931  CTGATACCTCCGACACACCTGCTTACTTCACTTCAATGGAGTGTAGCTGGCATCTCC 2990
Qy      971  AspAlaIleAsnAsnGlyTyzGlnSerTrpGlnSerTrpGlyProLeuPheGlyIleLeuPhePheAla 990
Db      2991  AATGCGATCAACACGATATGATCATGCGGCCCTGTTCGGAGCTCTTCTTTGCA 3050
Qy      991  PheTrpValIleValHisLeuTyzProPheLeuLeuLeuLeuMetGlyArgGlnAsnArg 1010
Db      3051  TTTTGGGTGATCGTCCATCTTTACCGTTCCTCAAGGGTCTGTGGAGGCGAGACAG 3110
Qy      1011  ThrProThrValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrp 1030
Db      3111  AGCCCAACGATTGTCATTGTGTGTCATCTCTCGTGGCTTCGATCTTCTCGCTGCTTGG 3170
Qy      1031  ValArgIleAspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIle 1050
Db      3171  GTCCCGATCGACCCGTTCTTGGCAAGGATGATGTCCTCTGTGGAGGATGTGTCTG 3230
Qy      1051  AsnCys 1052
Db      3231  GATTGC 3236
```

RESULT 4

```
US-10-160-719A-25
; Sequence 25, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Drugga, Karwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; PRIORITY FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 3813
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (215)...(3496)
US-10-160-719A-25
```

```
Alignment Scores:
Pred. No.: 0 Length: 3813
Score: 3974.50 Matches: 752
Percent Similarity: 78.53% Conservative: 122
Best Local Similarity: 67.57% Mismatches: 158
Query Match: 70.37% Indels: 81
DB: 4 Gaps: 17
```

US-10-627-132-30 (1-1052) x US-10-160-719A-25 (1-3813)

```
Qy      1  MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGlnLeuValLeuIle 20
Db      215  ATGGAGGCTAGCGCGGGCTGTGGCCGGCTCGCATNACCGAAGCTGTGTGTGTGTCGTC 274
```

```
Qy      21  ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnVal----- 36
Db      275  CGC-----CGCGACCGCGAGTCGGGAGCCGCGCGCGCGCGCGCGCGCGCGGAG 328
Qy      37  -----CyeGluIleCysGlyAspGluValGluThrValAspGlyAspLeuPheVal 54
Db      329  GCGCGCTGCCAGATATGCGCGACGAGTTCGGGGTGGGGCTTCGACGGGAGGCCCTTCGTG 388
Qy      55  AlaCysAsnGluCysGlyPheProValCysArgProCysTyzTrpGluThrGluArgArgGlu 74
Db      389  GCGTGCACGAGTGGCGCTTCCCGCTGCTCGCGCGCTGCTACGAGTACGAGCGCGCGAG 448
Qy      75  GlyThrGlnAsnCysProGlnCysLysThrArgTyzLysArgLeuLysGlySerProArg 94
Db      449  GCGTCGCAAGCGTGGCGCGAGTGCAGTGCAGCGCTCAAGCGCTCAAGGGCTCCCGCGCG 508
Qy      95  ValAlaGlyAspAspAspGluAspIleAspAspLeuGluHisGluPheAsnIleAsp 114
Db      509  GTGGCGCGGACGAGGAGGAGCGGCTGCGAGACCTGGAGGGCGAGTTCGGCCCTGCAG 568
Qy      115  AspGluAsnGlnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAla 134
Db      569  GACGCGCGCGCCACGAGGAC-----GACCCCGCAGTACGTCCCGCGAGTCC 613
Qy      135  MetLeuHisGlyArgMetSerTyzGlyArgGlyProAspAspGlyAspGly---AsnAsn 153
Db      614  ATGCTCAGCGCGCAGATGAGTACGCGCGCGCGCGCGCGCGCGCGCTTCAGCCCG 673
Qy      154  ThrProGlnIleProProIleIleThrGlySerArgSerValProValSerGlyGlu--- 172
Db      674  GTCCCCAAGTGGCGCTCTCCACCAACGCGCAGATGGTGTGATGACATCCCGCGCGAGCAG 733
Qy      173  -----PhePro-----IleThrAsnGlyTyzGlyHisGlyGluValSerSer 187
Db      734  CACGCGCTCGTCCCTCTACATGAGCGCGCGCGCGCGCGCGCGCG----- 778
Qy      188  LeuHisLysArgIleHisProTyzProValSerGluProGly-----SerAla 203
Db      779  -----AAGAGGATCCACCGCTCTTCGAGATCCCAACCTTCCAGTGCACCGCAGA 832
Qy      204  LysTrpAspGluLysGlu-----ValSerTrpLysGlu 215
Db      833  TCCATGGACCCGTCACAGGATCTGCGCCCTACCGATATGGCAGCTGCGCTGGAGGAG 892
Qy      216  ArgMetAspAspTrpLysSerLysGlnGlyIleLeu-----GlyGlyGly 230
Db      893  AGAATGGAGGGCTGGAAGCAGAGCAGGCGCGCTGCAGCATGTTCAGGAGCGAGGTGCG 952
Qy      231  AlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnProLeu 250
Db      953  GGTGATTGGGATGGCGAGCATCTGCCACTAATGGATGAAGCTAGGCGAGCATTTG 1012
Qy      251  SerArgLysValSerIleAlaSerSerLysValAsnProTyzArgMetValIleVal 270
Db      1013  TCCAGAAAAGTCCCTATATCATCAAGCCGAATTAATCCCTACAGATGATTAATCGTTATC 1072
Qy      271  ArgLeuValValLeuAlaPhePheLeuArgTyzArgIleLeuHisProValProAspAla 290
Db      1073  CGTTTGGTGGTTTGGGTTTCTTCTTCCATCCAGTACCGAGTGCATCCCGCGAAGATGCA 1132
Qy      291  IleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIleLeu 310
Db      1133  TTTGCAATTGTGGCTCATATCTGTAATCTGTGAAATCTGTTTTCGATGTCTCTGGATTCTT 1192
Qy      311  AspGlnPheProLysTrpPheProIleAspArgGluThrTyzLeuAspArgLeuSerLeu 330
Db      1193  GATCAGTTCCCAAGTGGCTTCCAAATCGAGAGAGAGACTTACTCGAGCGCTTGTGCATCA 1252
Qy      331  ArgTyzGluArgGlyGluProSerLeuLeuSerAlaValAspLeuPheValSerThr 350
Db      1253  AGTTTTCACAGGAGGTCACCCCTCTCAGCTTGTCTCCATCGACTTCTTGTCTAGTAGC 1312
Qy      351  ValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAla 370
```


RESULT 5

US-10-160-719A-45
 ; Sequence 45, Application US/10160719A
 ; Patent No. 6803498
 ; GENERAL INFORMATION:
 ; APPLICANT: Dhugga, Kanwarpal S.
 ; APPLICANT: Helentjaris, Timothy G.
 ; APPLICANT: Bowen, Benjamin A.
 ; APPLICANT: Wang, Xun
 ; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
 ; FILE REFERENCE: 0864C
 ; CURRENT APPLICATION NUMBER: US/10/160,719A
 ; PRIORITY FILING DATE: 2002-06-03
 ; PRIOR APPLICATION NUMBER: US 60/096,822
 ; PRIOR FILING DATE: 1998-08-17
 ; PRIOR APPLICATION NUMBER: US 09/371,383
 ; PRIOR FILING DATE: 1999-08-06
 ; NUMBER OF SEQ ID NOS: 61
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 45
 ; LENGTH: 3813
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (215)...(3496)
 ; US-10-160-719A-45

Alignment Scores:

Pred. No.: 0 Length: 3813
 Score: 3974.50 Matches: 752
 Percent Similarity: 78.53% Conservative: 122
 Best Local Similarity: 67.57% Mismatches: 158
 Query Match: 70.37% Indels: 81
 DB: 4 Gaps: 17

US-10-627-132-30 (1-1052) x US-10-160-719A-45 (1-3813)

Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuile 20
 Db 215 ATGGAGGTAGCGCGGGCTGGTGGCGGCTCGCATACCGGACGAGCTGGTGTGATC 274
 Qy 21 ArgGlyHisGluAspProGlyProLeuArgAlaLeuSerGlyGlnVal----- 36
 Db 275 CGC-----CGCAGCCGCGAGTCGGGAGCGCGCGCGCGCGCGCGCGCGCGAG 328
 Qy 37 -----CysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheVal 54
 Db 329 CGCGCGTCCAGATATGCGGCGAGCAGAGTCGGGTGGGCTTCGACGGGAGCCCTTCGTG 388
 Qy 55 AlaCysAsnGluCysGlyPheProValCysArgProCysTyrGluTyrGluArgArgGlu 74
 Db 389 CGGTGCAACAGATGGCGCTTCCCGCTCTGCGCGGCTCTACGAGTACGAGCGCGCGAG 448
 Qy 75 GlyThrGlnAsnCysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArg 94
 Db 449 GGCTCGCAAGCGTGCCTCGAGTCGAGACCGCTACAGCGCTCAAGGGCTCGCGCGG 508
 Qy 95 ValAlaGlyAspAspGluGluAspIleAspLeuGluHisGluPheAsnIleAsp 114
 Db 509 GTGGCGGCGCAGGAGGAGGAGCGGCTGACGACCTGGAGGGCGAGTTCGGCTCGCAG 568
 Qy 115 AspGluAsnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAla 134
 Db 569 GACGGCGCGCCCGCAGGAGC-----GACCCGCGAGTACGTCCGCGAGTCC 613
 Qy 135 MetLeuHisGlyArgMetSerTyrGlyArgGlyProAspAspGlyAspGly---AsnAsn 153
 Db 614 ATGCTCAGGGCGCAGATGAGTACGGCGCGCGCGCGCGCGCGCGCGCTTCAGCCCC 673
 Qy 154 ThrProGlnIleProProIleIleThrGlySerArgSerValProValSerGlyGlu--- 172

Db 674 GTCCCAACAGTGGCGGCTCCTCACCAACGGCCAGATGGTTGATGATCATCCCGCGGAGCAG 733
 Qy 173 -----PhePro-----IleThrAsnGlyTyrGlyHisGlyGluValSerSerSer 187
 Db 734 CACGGCTCGTGGCGTCTCTACATGAGCGCGCGCGCGCGCGCGCGCG----- 778
 Qy 188 LeuHisLysArgIleHisProTyrProValSerGluProGly-----SerAla 203
 Db 779 -----AAGAGGATCCACCGCTCCCTTTCGAGATCCCAACCTTCCAGTGCACCGAGA 832
 Qy 204 LysTrpAspGluLysLysGlu-----ValSerTrpLysGlu 215
 Db 833 TCCATGGACCGCTCAAGGATCTGGCCCGCTACGGATATGGCAGCTGGCTGGAGGAG 892
 Qy 216 ArgMetAspAspTrpLysSerLysGlnGlyIleLeu-----GlyGlyGly 230
 Db 893 AGAATGGAGGCTGGNAGCAGACGAGCGCGCTGCAGCATGTGCAGGCGGAGGTGGC 952
 Qy 231 AlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnProLeu 250
 Db 953 GGTGATTGGATGGCGAGCATCTGCCACTAATGATGAAGCTAGGCTAGGCGCATTG 1012
 Qy 251 SerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleVal 270
 Db 1013 TCCAGAAAGTCCCTATATCATCAAGCGCAATTATCCCTACAGGATGATTATCGTTATC 1072
 Qy 271 ArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAspAla 290
 Db 1073 CGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1132
 Qy 291 IleGlyLeuTrpLeuValSerIleCysGluIleTrpPheAlaIleSerTrpIleLeu 310
 Db 1133 TTTGCAATTGGCTCATATCTGTAATCTGTAATCTGTAATCTGTAATCTGTAATCTT 1192
 Qy 311 AspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSerLeu 330
 Db 1193 GATCAGTTCCCAAGTGGCTTCCCAATCGAGAGAGAGACTTACCTGGAGCGGTTTGTCACTA 1252
 Qy 331 ArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSerThr 350
 Db 1253 AGTTTTCAGCAGGAGGTCAACCTCTCAGCTTCTCCATCGACTCTTTTGTTCAGTAGC 1312
 Qy 351 ValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAla 370
 Db 1313 GTTGATCCCAAGAGAACTCCCTTGGTACAGCGAACAACCTGCTTCCATCTCTTCT 1372
 Qy 371 ValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLeu 390
 Db 1373 GTGGATTATCGGTGAGAGGCTCTCTGCTATGTTTCTGATGATGGTGTGCAATGCTT 1432
 Qy 391 ThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLys 410
 Db 1433 ACGTTTTCAGCATGTTCTGAAACATCTGAAATTCGAAAGAAATGGTTCCTTTCAGCAA 1492
 Qy 411 LysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeu 430
 Db 1493 AAGTTTAAATTCGAGCTCGTGGTCTCTGAGTGGTACTTCCAAACAGAGATAGACTACCTG 1552
 Qy 431 LysAspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGluTyrGlu 450
 Db 1553 AAAGCAACAGGTGCTGCTTCAATTTGAGGAGGAGGCGGATGAAGAGAGAAATACCGAG 1612
 Qy 451 GluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGluGly 470
 Db 1613 GAATTCAGGTAGGATCAATGCTTGGTTCGAAAGCCCAAGAGGTTCCTGTGAGAGAGA 1672
 Qy 471 TrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGlyMet 490
 Db 1673 TGGCAATGCAAGATGGAAGCCCTGGCTGGAAACAACAGTACGCGATCATCTCTGGAATG 1732
 Qy 491 IleGlnValPheLeuGlyHisSerGlyHisAspThrGluGlyAsnGluLeuProArg 510

```
Db 1733 ATTCAGGTATTCCTTGCCAAAGTGGCGTCTGATGTGGAAGAAATGAGTTGGCTCGC 1792
Qy 511 LeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGlyAla 530
Db 1793 CTGGTTTATGTCTCGAGAGAAAGAGCCAGGTTATAACCATCAACAAGAGGCTGGTGCC 1852
Qy 531 MetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsnLeu 550
Db 1853 ATGAAATGACATGGTCCGCTGCTCTGCTGTCTTATCAAAATGCTGCATACCTATTGAACTTG 1912
Qy 551 AspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeuMet 570
Db 1913 GACTGTGATCACTACATCAACAATAGCAAGGCCATAAAGAGGCTATGTGTTTCATGATG 1972
Qy 571 AspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIle 590
Db 1973 GATCCTTTGGTGGGGAAGAAAGTGTGTATGTATGACAGTTCCCTCAGAGGTTTGATGGTATT 2032
Qy 591 AspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLysGly 610
Db 2033 GACAAAATGATCATACGTATACAGGAACGTTGCTTTTTTGACATCAACATGAAGGT 2092
Qy 611 LeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGlnAla 630
Db 2093 TTGGACGGTATTCAAGGACCCATTATGTGGGTACTGGATGTGTTTTCAGACGGCAGGCA 2152
Qy 631 LeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysAspCys 650
Db 2153 CTGTATGTTATGATGCTCTCTAAA---ACGAAGAAGCCACCATCAAGAACTTGCACACTGC 2209
Qy 651 CysPro-----CysPheGlyArgLysLysArgLys----- 660
Db 2210 TGGCCCAAGTGGTGCCTCTCTGTCTGTGACAGGAAACAAGATAAAGAAAGACTACA 2269
Qy 660 ----- 660
Db 2270 AAACCAAGACGGAGNAGAAAGAAAGATTATTTTCAAGAAAGCAGAAACCCATCTCCT 2329
Qy 661 ---HisAlaLysAspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAspLys 679
Db 2330 GCATATGCTTTGGGTGAAATTTGATCAAGGTCTCTCA-----GGTGTGATATCGAGAAG 2383
Qy 680 GluMetLeuSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheVal 699
Db 2384 GCGGAATCGTAATCAACAGAACTAGAGAAGAAATTTGGCGAGTCTCTGTTTGTCT 2443
Qy 700 ThrSerThrLeuMetGluGluGlyValProProSerSerSerProAlaAlaLeuLeu 719
Db 2444 GCATCAACACTTCTTGAGAACGGAGGACCCTGAAGAGCGCAAGTCCACAGCTTCTCTCTG 2503
Qy 720 LysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTyrGlyLeuGlu 739
Db 2504 AAGGAAGCTATACATGTTATCAGCTGCGGCTACGAAGAACAAGCCGACTGGGGAAGAG 2563
Qy 740 LeuGlyTyrPheTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCys 759
Db 2564 ATTGCTGGAATTTACGGATCGATCAAGAGATATCTTGATGGATTAAAGTCACACTGC 2623
Qy 760 ArgGlyTyrArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaPro 779
Db 2624 CATGCTGGCGGTCTATTACTGATCGCGCTCGAAGCGGCTGCATTCAAAGGTTCTGCGCT 2683
Qy 780 IleAsnLeuSerAspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIle 799
Db 2684 CTGAACCTTTCCGACCGCTTTCACAGGTCCTTCGCTGGGCGCTTGGGTCCGTGAAAT 2743
Qy 800 PhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTyrLeu 819
Db 2744 TTCCTCAGACAGCATGCGCCACTTTGGTACGGATAC---GGCGGGCGGCTAAATTCCTG 2800
Qy 820 GluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAla 839
Db 2801 GAAAGGTTTCTTATATCAACTCCATCGTTTATCGCTGAGACGTCATTCCTCTCTCGCT 2860
```

```
Qy 840 TyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSer 859
Db 2861 TACTGTACCTTGCTGCCATCTGCTGCTCAACGGGAGTTTATCACACAGAGCTTACC 2920
Qy 860 ThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeu 879
Db 2921 AATGTCCGCAATATCTGGTTTCATGGCACTTTTCATCTGCATCTCCGTGACCGCATCCTG 2980
Qy 880 GluMetArgTyrSerGlyValSerIleGluGluTyrTyrArgAsnGluGlnPheTyrVal 899
Db 2981 GAAATCAGGTGGAGTGGCGCATCGACACTGTGTGGAGGAACGAGCAGATTCTGGGTC 3040
Qy 900 IleGlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAla 919
Db 3041 ATCGAGGGCTTTGGGGCATCTGTTCGGGTCTTCAGGGCTCTGCTGAAGGTGTTCCGCC 3100
Qy 920 GlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPhe 939
Db 3101 GGCATCGACACGAGCTTCACCGTGACGTGCAAGGCC---GGGGAC---GACGAGGAGTTC 3154
Qy 940 AlsGluLeuTyrAlaPheLysTyrThrLeuLeuIleProProThrThrLeuLeuIle 959
Db 3155 TCGAGAGCTGTACACGTTCAAGTGGACACCCCTGCTGATACCCCGACCCGCTCCTCTG 3214
Qy 960 IleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSer 979
Db 3215 CTGAACTTCATCGGGTGGTGGCGGGATCTCGAACCGCATCAACAACGGGTACGAGTCG 3274
Qy 980 TrrGlyProLeuPheGlyLysLeuPhePheAlaPheTyrValIleValHisLeuTyrPro 999
Db 3275 TGGGGCCCCCTGTTCGGGAAGCTCTTCTCGCCTTCTGGGTGATCGTCCACCTGTATCCCG 3334
Qy 1000 PheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIleTrrSer 1019
Db 3335 TTCTCTCAAGGTCTGTGGGAGGAGAGAACAGGACGCGCAGATCGTCTGCTGTGTCTCC 3394
Qy 1020 IleLeuLeuAlaSerIlePheSerLeuLeuTyrValArgIleAspProPheIleValArg 1039
Db 3395 ATCTGCTGGCTCGATCTTCTGCTCTGTGGGTCCGCGTCCGACCCGCTTCTCGCCAAG 3454
Qy 1040 ThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3455 AGCAACGGCCGCTCTCTGGAGAGTGTGGCTGGAGTGC 3493
```

RESULT 6

```
US-10-160-719A-29
; Sequence 29, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 3746
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (321)...(3551)
US-10-160-719A-29
```

Alignment Scores:
 Pred. No.: 0 Length: 3746
 Score: 3912.50 Matches: 750
 Percent Similarity: 78.96% Conservative: 102
 Best Local Similarity: 69.51% Mismatches: 150
 Query Match: 69.27% Indels: 77
 DB: 4 Gaps: 17

US-10-627-132-30 (1-1052) x US-10-160-719A-29 (1-3746)

Qy	34	GlyGlnValCysGluileCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPhe	53
Db	366	GGACAGGTGTCAGATCTGGCGACGGCTGGGCACACGGCGGAGGGAGCGTCCTTC	425
Qy	54	VallalaCysAenGluCysGlyPheProValCysArgProCysTyrGluTyrGluArgArg	73
Db	426	GCCGCTCGCAGCGTCTGGGGTTCCTCGGTGTGCGCCCTGCTACGAGTACGAGCGCAAG	485
Qy	74	GluGlyThrGlnAenCysProGlnCysIysThrArgTyrLysArgLeuLysGlySerPro	93
Db	486	GACGCGACGACGGCTGCCCGAGTGAAGACCAAGTACAAAGCGCCACAAAGGGAGCCGC	545
Qy	94	ArgValAlaGlyAspAspGluGluAspIleAspLeuGluHisGlyPheAsn--	112
Db	546	CGCATCCGTGGG-----GAGNAGGACGACACTGATCGCATAGCGACTTCAATTAC	599
Qy	113	IleAspAspGluAenGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThr	132
Db	600	CTTGCACTCGCAATAGGAC-----CAGAAGCAGAGAGATTGCC	638
Qy	133	GluAlaMetLeuHisGlyArgMetSerTyrGlyArgGlyProAspAspGly--	149
Db	639	GACAGAATGCCAGCTGGCGCATGAACCTTGGGGGACGCGGGGATGTTGGTCGCCCAAG	698
Qy	150	-----AepGlyAsnAsnThrProGln--Ile	157
Db	699	TATCACAGTGGCGAGATCGGCTTACCAAGTATGACAGTGGCGAGATTCCTCGGGGATAC	758
Qy	158	ProProlleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsn	177
Db	759	ATCCCATCAGTCACTAAACAGCGAG-----ATCTCAGGAGAAATCCCTGGTGCTTCC	809
Qy	178	GlyTyrGlyHis--GlyGluValSerSerLeuHisLysArgIle--HisProTyr	195
Db	810	CCTGACCATCATATGATGTCCCAACTGGGAACATTTGGCAAGCGTGCTCCATTTCCCTAT	869
Qy	196	ProValSerGluProGlySerAlaLysTrpAspGluLysLys-----GluValSerTrp	213
Db	870	GTGAACCATTCGCCAAATCCGTCAAGGAGTCTCTGTGTAGCATTTGGGAATGTTGCCTGG	929
Qy	214	LysGluArgMetAspAspTrpLysSerLysGln-----GlyIleLeu-----	227
Db	930	AAAGAGGGTGTGATGGCTGGAAATGAAGCAGGACAAAGGGGCGCATTTCCCATGACGAA	989
Qy	228	-----GlyGlyGlyAlaAspProGluAspMetAspAla-----	238
Db	990	GGCACAGCATTTGCTCCCTCTGAGGTTCGGGTGTGTGTATATTGATGCATCACTGAT	1049
Qy	239	-----AspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysVal	254
Db	1050	TACAAACATGGAAGATGCTTATTGAACGACGAAACTCGACAGCCTCTATCTAGGAAAGTT	1109
Qy	255	SerIleAlaSerSerLysValAsnProTyrArgMetValIleValValArgLeuValVal	274
Db	1110	CCATCTCCTTCCTCAGATAAATCCATACAGGATGTCATTTGCTGCGCATGATGATT	1169
Qy	275	LeuAlaPheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrp	294
Db	1170	CTAAGCATCTCTTGCACTACCGTATCACAAATCCTGTGGCATGATGATGATGATG	1229
Qy	295	LeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPhePro	314

```
QY 662 AlAlAspGlyLeuPro-----GluglyThrAlaAspMet 673
Db 2361 GTGACAGTTCTGTGCCAGTATTCAACCTTGAAGATATAGAGGAGGAGTTGAAGCGCT 2420
QY 674 GlyValAspSerAspGlyLeuMetLeuMetSerHisMetAenPheGluLysArgPheGly 693
Db 2421 GGATTGACGACGAGAAATCATTCTTATGTTCTCAATGAGCCTGGAGAGAGATTGGC 2480
QY 694 GluSerAlaAlaPheValThrSerThrLeuMetGluGluGlyValProProSerSer 713
Db 2481 CAGTCCGACGCTTTGTTCCTCCACTCTGATGAGTATGTTGTGTTCTCTCAGTCGCA 2540
QY 714 SerProAlaAlaLeuLeuLysGluAlaHisValIleSerCysGlyTyrGluAspLys 733
Db 2541 ACTCCGAGTCTCTTCTGAAGAGACTATCCATGTTATAGCTGTGGCTATGAGCAAG 2600
QY 734 ThrAspTrpGlyLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThr 753
Db 2601 ACTGAATGGGAACTGAGATCGGTGGATCTACCGTTCTGTGACAGAGACATCTCACC 2660
QY 754 GlyPheLysMetHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAla 773
Db 2661 GGATTCAAGATGACGCGCGAGGCTGCGGTGCTACTTGCATGCCCAAGCGGCCAGCT 2720
QY 774 PheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAla 793
Db 2721 TTCAAGGGGTCTGCCCATCATCATCTTTTCGACCGCTCTGAACCAAGGTCCTCGGTGGCT 2780
QY 794 LeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsn 813
Db 2781 CTTGGGTCGCGGAGATCTCTTTGACCGGCACCTGCCCCCTGTGTACGGGTAC--GGA 2837
QY 814 GlyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrIleTyrProPheThr 833
Db 2838 GSGCGGCTCAAGTTCTCTGAGAGATTGCGGTACATCAACACCATCTACCCGCTCAGC 2897
QY 834 SerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPhe 853
Db 2898 TCCATCCCGCTTCTCATCTACTGATCCGCGGCATCTGTGCTGCTACCGGAAAGTTC 2957
QY 854 IleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIle 873
Db 2958 ATCATTCAGAGATCAGCAACTTCCGACGATCTGGTTTCATCTCCCTTTCATCTCGATC 3017
QY 874 PheAlaThrGlyLeuLeuMetArgTrpSerGlyValSerIleGluGluTrpTrpArg 893
Db 3018 TTGCGCACGGGCATCTCGAGATGAGTGGAGCGGGTGGGCATCGACAGTGTGGAGG 3077
QY 894 AsnGluGlnPheTrpValIleGlyValSerAlaHisLeuPheAlaValGlnGly 913
Db 3078 AACGAGCAGTTCTGGGTGATCGGGGGCATCTCCGCGCACCTCTTCCGCGTTTCCAGGGC 3137
QY 914 LeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGly 933
Db 3138 CTGCTCAAGGTCTGGCGGCATCGACACCACTTCCACCGTCACTCCAGGCGCTCG-- 3194
QY 934 AspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIlePro 953
Db 3195 GACGAGGACGGGACTTCCGGGAGCTGTACATGTTCAAGTGGACGACGCTCTCGATCCG 3254
QY 954 ProThrThrLeuLeuIleAsnValIleGlyValValAlaGlyIleSerAspAlaIle 973
Db 3255 CCCACCATCTCTGATCATCAACTGTCGGCGCTCGTCGGCGGATCTCTTACGCCATC 3314
QY 974 AsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpVal 993
Db 3315 AACAGGGATACAGTGTGGGGCCGCTCTTCGCAAGCTCTTCTTCGCTTCTGGGTC 3374
QY 994 IleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThr 1013
Db 3375 ATCGTCCACCTGTACCCGTTCTCAAGGGCCTCATGGGCGGCTCATGGGCGAGACCGCACCCCGACC 3434
```

```
QY 1014 ValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIle 1033
Db 3435 ATCGTGTGTCTGGGCCATCTCTGTGGCTCCATCTTCTCTTGTGTGGTTCGCATC 3494
QY 1034 AspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3495 GACCCCTTACCACCCGCGTCTCTGTCGCGGATACCCAGACGTTGGCATCAACTGC 3551

RESULT 7
US-10-160-719A-49
; Sequence 49, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 3746
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (321) ... (3449)
US-10-160-719A-49

Alignment Scores:
Pred. No.: 0 Length: 3746
Score: 3912.50 Matches: 750
Percent Similarity: 78.96% Conservative: 102
Best Local Similarity: 69.51% Mismatches: 150
Query Match: 69.27% Indels: 77
DB: 4 Gaps: 17

US-10-627-132-30 (1-1052) x US-10-160-719A-49 (1-3746)
QY 34 GlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPhe 53
Db 366 GGACAGGTGTGCCAGATCTCGCGCACGGCTGGGCACCCACGCGGAGGGAGCGTCTTC 425
QY 54 ValAlaCysAsnGluCysGlyPheProValCysArgProCysTrpGluTyrGluArgArg 73
Db 426 GCCGCTTCGACGCTCTCGGGTTTCGCGCTTCCGCTGTCGCCCTGCTACGAGTACGAGCGCAAG 485
QY 74 GluGlyThrGlnAsnCysProGlnCysLysThrArgTyrLysArgLeuLysGlySerPro 93
Db 486 GACGCGACGACAGGCGTCCCTCCAGTGCAGACCAAGTACAGCCGCCACACAGGGGAGCGCG 545
QY 94 ArgValAlaGlyAspAspAspGluAlaAspIleAspAspLeuGluHisGluPheAsn--- 112
Db 546 GCGATCCGTGGG-----GAGGAAGGAGACGACACTGATGCCGATAGCGACTTCAATTAC 599
QY 113 IleAspAspGluAsnGlnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThr 132
Db 600 CTTGCATCTGCAATGAGGAC-----CAGAACGACGAGAGATTGCC 638
QY 133 GluAlaMetLeuHisGlyArgMetSerTyrGlyArgGlyProAspAspGly----- 149
Db 639 GACAGATGCCACTGGCGCATCAACGTTGGGGGACGCGGGGATGTTGGTCGCCCCCAAG 698
QY 150 -----AspGlyAsnAsnThrProGln-----Ile 157
```

Db 699 TATGACAGTGGCGAGATCGGCTTACCAAGTATGACAGTGGCGGAGATTCCTCGGGGATAC 758
Qy 158 ProProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAen 177
Db 759 ATCCCATCATGCTCACTAACAGCCAG-----ATCTCAGGAGAAATCCCTGGTCTCC 809
Qy 178 GlyTyrGlyHis---GlyGluValSerSerSerLeuHisLysArgIle---HisProTyr 195
Db 810 CTGACCATCATATGATGTCCTCCCAACTGGGAACATGGCAAGCGTCTCCATTCCTCCAT 869
Qy 196 ProValSerGluProGlySerAlaLysTrpAspGluLysLys-----GluValSerTrp 213
Db 870 GTGACCATTCGCCAAATCCGTCAAGGAGTTCCTCTGGTAGCATTCGGGAATGTCCTCG 929
Qy 214 LysGluArgMetAspAspTrpLysSerLysGln-----GlyIleLeu----- 227
Db 930 AAAGAGAGGGTTCATGGCTGGGAATGAAGCAGGACCAAGGCGACGATCCCATGACGAAT 989
Qy 228 -----GlyGlyAlaAspProGluAspMetAspAla----- 238
Db 990 GGCACAAAGCATTCCTCCCTCTGAGGTCGGGTGGTGGTATGATGTCATCAACTGAT 1049
Qy 239 -----AspValAlaLeuAenAspGluAlaArgGlnProLeuSerArgLysVal 254
Db 1050 TACAACATGGAAGATGCCTTATTGACGACGAACTCGACAGCCCTCTATCTAGGAAGTT 1109
Qy 255 SerIleAlaSerSerLysValAenProTyrArgMetValIleValValArgLeuValVal 274
Db 1110 CCATTCCTCTCTCCAGGATAAATCCATACAGGATGGTTCATTGTGCTGCGATTGATTGT 1169
Qy 275 LeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrp 294
Db 1170 CTAAAGCATTCCTTCGACCTACCGTATCAAAATCCTGTGCGCAATGCATACCCATTATGG 1229
Qy 295 LeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPhePro 314
Db 1230 CTTCATCTGTATATGAGATCTGTTTGTCTCTTTCTGTGATATTTGATCAGTTCCT 1289
Qy 315 LysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArg 334
Db 1290 AAGTGGTTTCCAATCAACCGGAGACGTACCTTTGATAGCTGGCATTAAGGTATGACCGG 1349
Qy 335 GluGlyGluProSerLeuSerAlaValAspLeuPheValSerThrValAspProLeu 354
Db 1350 GAAGGTGAGCCATCTCAGTTGGCTGCTGTGACATTTTTCGTGACGTACAGTCGACCAATG 1409
Qy 355 LysGluProProLeuValThrAlaAenThrValLeuSerIleLeuAlaValAspTyrPro 374
Db 1410 AAGGAGCTCTCTTGTCTACTGCCAATACCGTCTATCCATTTCTTGTGCTGTGGATTACCCT 1469
Qy 375 ValAspLysValSerCysTyrValSerAspGlyAlaSerMetLeuThrPheGluSer 394
Db 1470 GTGGATAAGTCTCTTGTATGTATCTGATGATGAGCTGCGATCTGCATTTTGATGCA 1529
Qy 395 LeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIle 414
Db 1530 CTAGCTGAGACTTTCAGATTGCTAGAAAATGGGTACCATTTTGTAAAGAAGTACAAACAT 1589
Qy 415 GluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysVal 434
Db 1590 GAACCTAGAGCTCTGAAATGGTACTTCTCCGAAAATTTGATTACTTGAAGGACAAAGTG 1649
Qy 435 GlnProThrPheValGlnGluArgAlaMetLysArgGluTyrGluGluPheLysVal 454
Db 1650 CACCTTCATTTGTTAAAGACCGCGCGCCATGAAGAGAGAATAATGAAGAATTCAAAGTT 1709
Qy 455 ArgIleAenAlaLeuValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLys 474
Db 1710 AGGGTAATGGCTTGTGCTTAAGGACAGAAAAGTTCTCGAAGGAAGGATGATCATGCAA 1769
Qy 475 AspGlyThrProTrpProGlyAenAenThrArgAspHisProGlyMetIleGlnValPhe 494
Db 1770 GATGSCACACCATGGCCAGGAACAAATACCAGGGACCATCTCTGGAAATGATTTCAGGTTTC 1829

Qy 495 LeuGlyHisSerGlyGlyHisAspThrGluGlyAenGluLeuProArgLeuValTyrVal 514
Db 1830 CTTGTGTACAGTGGTGGCTTGATCTAGGGCAATAGTACCTACCGCTTTGTCTATGTT 1889
Qy 515 SerArgGluLysArgProGlyPheGlnHisIshisLysLysAlaGlyAlaMetAenAlaLeu 534
Db 1890 TCTCGTGAAGAGCGTCTGGATTCAGCATCAAGAAGAGCTGGTGCCATGAATGCTCTT 1949
Qy 535 IleArgValSerAlaValLeuThrAenAlaProPheMetLeuAenLeuAspCysAspHis 554
Db 1950 GTTCGTGTCTCAGCTGTCTTACCATGACATACATGTTGAATCTTGATTTGTATCATC 2009
Qy 555 TyrIleAenAenSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnVal 574
Db 2010 TACATTAACAACAGTAAGGCTCTCAGGGAAGCATGTGCTCTCTTATGACCTAACCTA 2069
Qy 575 GlyArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAsp 594
Db 2070 GGAAGGAGTGTCTGTCTACGTCCAGTTTCCAGAGATTCGATGGCATTCGACAGGAATGAT 2129
Qy 595 ArgTyrAlaAenArgAsnThrValPhePheAspIleAenMetLysGlyLeuAspGlyIle 614
Db 2130 CGATATGCCACAGACAGACACCGTGTCTTTCATATTAATCTTGAGAGGTCTTGATGGCATC 2189
Qy 615 GlnGlyProValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyr 634
Db 2190 CAAGACCACTTATGTCCGAACCTGGCTGTGTTTTCAACCGCAACAGCTCTATATGTTAT 2249
Qy 635 AsnProProLysGlyProLysArgProLysMetValThrCysAspCysProCysPhe 654
Db 2250 GAGCCCCCAATTAAGACAGAGAGGGTGTGTTCTTGTCA-----TCACTATGTGCG 2300
Qy 655 GlyArgLysLys-----ArgLysHis 661
Db 2301 GGTAGGAAGAGGACCAAGCAATCAAGAGGGCTCGGACAAAGCAAGTCCGCAAGGAT 2360
Qy 662 AlaLysAspGlyLeuPro-----GluGlyThrAlaAspMet 673
Db 2361 GTGGACAGTCTGTGTCAGTATTCAACCTTGAAGATATAGAGGAGGGAGTTGAAGCGCT 2420
Qy 674 GlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGly 693
Db 2421 GGATTTGACGACGAGAAATCACTTCTTATGTCTCAAAATGAGCCTGGAGAGAGATTTCGC 2480
Qy 694 GlnSerAlaAlaPheValThrSerThrLeuMetGluGlyGlyValProProSerSer 713
Db 2481 CAGTCCGCGAGGTTTGTGCTCCACTCTGATGAGTATGTTGTTGTTCTTCAGTCCGCA 2540
Qy 714 SerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLys 733
Db 2541 ACTCCGAGTCTCTTCTGAAGAAGTATCCATGTTATTAAGCTGTGGCTATGAGGACAAG 2600
Qy 734 ThrAspTrpGlyLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThr 753
Db 2601 ACTGAATGGGAACTGAGATCGGTGGATCTACCGTCTCTGTGACAGAGAATTCCTCACC 2660
Qy 754 GlyPheLysMetHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAla 773
Db 2661 GGATTCAGATGACGCGCGAGGCTGGCGTCTGATCTACTGATGCTCCCAAGCGGCAGCT 2720
Qy 774 PheLysGlySerAlaProIleAsnLeuSerAspArgLeuAenGlnValLeuArgTrpAla 793
Db 2721 TTCAAGGGGTCTGCCCCCATCAATCTTTCGACCGTCTGAACACAGGTCTCCGGTGGCT 2780
Qy 794 LeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAen 813
Db 2781 CTTGGGTCCGCGGAGATCTCTCTCAGCGGCACCTGCCCCCTGTGTACGGTAC---GGA 2837
Qy 814 GlyAenLeuLysTrpLeuGluArgPheAlaTyrIleAenThrIleTyrProPheThr 833
Db 2838 GGGCGGCTCAAGTCTCTGGAGAGATTCCGTACATCAACACCATCTACCCGCTCAGG 2897

```
Qy 834 SerLeuProLeuLeuAlaTyrcysThrLeuProAlaValCysLeuLeuThrGlyLysPhe 853
Db 2898 TCCATCCGGTCTCTACTGATCTGCGCCGCATCTGCTGCTCACGGANAGTTC 2957
Qy 854 IleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIle 873
Db 2958 ATCATTCCAGAGATCAGCAACTTCGCCGAGCATCTGGTTTCATCTCCCTCTTCATCTCGATC 3017
Qy 874 PheAlaThrGlyIleLeuLeuMetArgTrpSerGlyValSerIleGluGluTrpTrpArg 893
Db 3018 TTCGCCACGGGCATCTCTGGAGATGAGTGGAGCGGGTGGGCATCGACGAGTGGTGAGG 3077
Qy 894 AsnGluGlnPheTrpValIleGlyValSerAlaHisLeuPheAlaValValGlnGly 913
Db 3078 AACGAGGAGTCTGGGTGATCGGGGGCATCTCCGGGCACCTCTTCGGCGGTTCAGGGC 3137
Qy 914 LeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGly 933
Db 3138 CTGCTCAAGGTGTCGGCGGCATCGACACCAACTTCACCGTCACTCCAGGGCCTCG--- 3194
Qy 934 AspGluAspAspGluPheAlaGluLeuTyrlalaphelysTrpThrThrLeuLeuIlePro 953
Db 3195 GACGAGGACGGGACTCTCGCGAGCTGTACATGTTCAGTGGACGACGCTCCTGATCCCG 3254
Qy 954 ProThrThrLeuLeuIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIle 973
Db 3255 CCCACCACTCTGATCATCAACTGCTGGTGGCGTCTGCGCGGCATCTCTACGCCATC 3314
Qy 974 AsnAsnGlyTyrlGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpVal 993
Db 3315 AACAGCGGATACCAAGTCTGGGGCCGCTCTTCGCGCAAGCTCTTCTTCGCGCTTCGGGTC 3374
Qy 994 IleValHisLeuTyrlProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThr 1013
Db 3375 ATCGTCCACCTGTACCCGCTTCTCAAGGGCTCTATGGCGAGCAGACCGCACCCCGACC 3434
Qy 1014 ValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIle 1033
Db 3435 ATCGTGTGCTGTGGGCATCTGCTGGCGTCCATCTCTCTGCTGTGGTTCGCATC 3494
Qy 1034 AspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3495 GACCCCTTCACCAACCCGGCTCACTGGCCCGGATACCCAGACGCTGGGCATCAACTGC 3551

RESULT 8
US-10-160-719A-5
; Sequence 5, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-08-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3773
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (338)...(3568)
US-10-160-719A-5
```

```
Alignment Scores:
Pred. No.: 0 Length: 3773
Score: 3912.50 Matches: 750
Percent Similarity: 78.96% Conservative: 102
Best Local Similarity: 69.51% Mismatches: 150
Query Match: 69.27% Indels: 77
DB: 4 Gaps: 17

US-10-627-132-30 (1-1052) x US-10-160-719A-5 (1-3773)
Qy 34 GlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPhe 53
Db 383 GGACAGTGTGCCAGATCTCGCGCGACGGCGTGGGCCACCGCGCGGGGGACGCTTC 442
Qy 54 ValAlaCysAsnGlnCysGlyPheProValCysArgProCysTyrlGluTyrlGluArg 73
Db 443 GCCGCTCGACGCTCTCGGGTTCGGGTGTGGCGCCCTGCTACGAGTACGAGCGCAAG 502
Qy 74 GluGlyThrGlnAsnCysProGlnCysLysThrArgTyrlLysArgLeuLysGlySerPro 93
Db 503 GACGGCACCGCAGCGCTGCCCGCAGTGCACAGCAAGTACAAAGCGCCACAAAGGGGAGCCG 562
Qy 94 ArgValAlaGlyAspAspAspGluGluAspIleAspAspLeuGluHisGluPheAsn--- 112
Db 563 GCGATCCGTGGG-----GAGGAAGGAGACACACTGATGCCGATACGCACTTCAATTAC 616
Qy 113 IleAspAspGluAsnGlnGlnArgGlnLeuGluGlnAsnMetGlnAsnSerGlnIleThr 132
Db 617 CTGTCATCTGGCATGAGGAC-----CAGAAGCAGAGAGATTGCC 655
Qy 133 GluAlaMetLeuHisGlyArgMetSerTyrlArgGlyProAspAspGly----- 149
Db 656 GACAGAATGCGCAGTGGCGCATGAACGTTGGGGCAGCGGGGATGTTGTGTCGCCCAAG 715
Qy 150 -----AspGlyAsnAsnThrProGln-----Ile 157
Db 716 TATGACAGTGGCAGATCGGGCTTACCAAGTATGACAGTGGCGAGATCTCTCGGGGATAC 775
Qy 158 ProProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsn 177
Db 776 ATCCATCATGCTACTAACGCCAG-----ATCTCAGAGANAATCCCTGGTGTCTCC 826
Qy 178 GlyTyrlGlyHis---GlyGluValSerSerSerLeuHisLysArgIle---HisProTyrl 195
Db 827 CCTGACCATCATATGATGTCCCAACTGGGAACATTTGGCAAGCGTGTCTTCCCTAT 886
Qy 196 ProValSerGluProGlySerAlaLysTrpAspGluLysLys-----GluValSerTrp 213
Db 887 GTGAACCATTCGCCAAATCCGTCAAGGGAGTCTCTGCTAGCATTTGGGAATGTTGCTGG 946
Qy 214 LysGluArgMetAspAspTrpLysSerLysGln-----GlyIleLeu----- 227
Db 947 AAAGAGAGGTGATGCTCGAANAATGACGAGCAGCAAGGGGACGATTCCCATGACGAAT 1006
Qy 228 -----GlyGlyAlaAspProGluAspMetAspAla----- 238
Db 1007 GGCACAAGCATTTGCTCCTCTGAGGGTCCGGGTGTTGTTGATATGATGATCACTAATGAT 1066
Qy 239 -----AspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysVal 254
Db 1067 TACAACATGGAAGATGCTTATTGAACGACGAAACTCGACAGCCTCTATCTAGGAAGTT 1126
Qy 255 SerIleAlaSerSerLysValAsnProTyrlArgMetValIleValValArgLeuValVal 274
Db 1127 CCACTTCTTCTCCAGGATAAATCCATACAGGATGGTCAITTTGTGTCGATTTGATTGT 1186
Qy 275 LeuAlaPhePheLeuArgTyrlArgIleLeuHisProValProAspAlaIleGlyLeuTrp 294
Db 1187 CTAAGCATCTTCTTGGCACTACCGTATCAAAATCCTGTGCGCAATGCAATGCCATTTATGG 1246
Qy 295 LeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPhePro 314
Db 1247 CTTCTATCTGTATATGTGAGATCTGGTTTCTCTTCGUGGATATTTGGATCAGTTCCT 1306
```


QY 315 LysTrpPheProIleAspArgGluThrThrLeuAspArgLeuSerLeuArgTyrGluArg 334
DB 1307 AAGTGGTTCCTCAATCAACGGGAGACGTACTCTTGATGAGCTGGCATTAAGGTATGACCGG 1366
QY 335 GluGlyGluProSerLeuSerAlaValAspLeuPheValSerThrValAspProLeu 354
DB 1367 GAAGGTGACCAATCTCAGTTGGCTGCTGTTGACATTTTCGTGAGTACAGTCGACCAATG 1426
QY 355 LysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrPro 374
DB 1427 AAGGAGCTCTCTTGTCACATGCCAATACCGTGCTATCCATTTCTTGCTGTGATTAACCT 1486
QY 375 ValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheGluSer 394
DB 1487 GTGGATTAAGTCTCTTGCTATGTATCTGATGATGGAGCTGCATGTGATTCGATCA 1546
QY 395 LeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIle 414
DB 1547 CTAGCTGAGACTTCAGAGTTTCTAGAAATGGTACCATTGTTAAGAAAGTACAACATT 1606
QY 415 GluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysVal 434
DB 1607 GAACCTAGAGCTCTGAAATGGTACTTCCACAGAAATTTGATTACTTGAAGGACAAAGTG 1666
QY 435 GlnProThrPheValGlnGluArgAlaMetLysArgGluTyrGluGluPheLysVal 454
DB 1667 CACCCTTCATTTTAAAGACCGCGCGGCATGAAGAGAGAGATATCAAGAATTCAAAGTT 1726
QY 455 ArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLys 474
DB 1727 AGGGTAATGGCTTGTGCTAAGGCACAGAAAGTTCTTGAGGAAGGATGGATCAATGCAA 1786
QY 475 AspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPhe 494
DB 1787 GATGCACACCATGCGCAGGAACAATACACAGGACCATCTCTGGAAATGATTCAGGTTC 1846
QY 495 LeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrVal 514
DB 1847 CTTGGTCAAGTGGTGGCTTGATCTAGGGCAATGAGTACCCCGCTTGGTCTATGTT 1906
QY 515 SerArgGluLysArgProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAlaLeu 534
DB 1907 TCTGTGAAAGCGTCTGGATTCAGACATCAAGAAAGCTGGTGCCATGAATGCTCTT 1966
QY 535 IleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHis 554
DB 1967 GTTCGTGCTCAGCTGCTTACCAATGGACATACATGTGAACTTGTATGTGATCAC 2026
QY 555 TyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnVal 574
DB 2027 TACATTAACAACAGTAAGCTCTCAGGGAAGCTATGTGCTTCTTATGGACCCCTAACCTA 2086
QY 575 GlyArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAsp 594
DB 2087 GGAAGAGTGTCTGTCTAGCTCAGTTTCCCCAGAGATTGATGGCAATTCAGACAGATGAT 2146
QY 595 ArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIle 614
DB 2147 CGATATGCCACAGGAACACCGTGTTCGATATTAATCTTGAGAGGCTTGTATGGCATC 2206
QY 615 GlnGlyProValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyr 634
DB 2207 CAAGGACCAATTTATGTCGGAATCGCTGTGTTTTCAACCGAAACAGCTCTATATGTTAT 2266
QY 635 AsnProProLysGlyProLysArgProLysMetValThrCysAspCysProCysPhe 654
DB 2267 GAGCCCCCAATTAAGCAGAAAGGGTGTGTTCTTTGTCA-----TCACTATGTGGC 2317
QY 655 GlyArgLysLys-----ArgLysHis 661
DB 2318 GGTAGAGAGAGGCAAGCAATCAAGAGGGCTCGGACAGAAAGATCGCAGAGCAT 2377

QY 662 AlaLysAspGlyLeuPro-----GluGlyThrAlaAspMet 673
DB 2378 GTGCAGACTTCTGTCCAGATATTCAACCTTGAAGATATAGAGGAGGATTTCAAGCGCT 2437
QY 674 GlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGly 693
DB 2438 GGATTTGACGAGGAATCTACTTCTTATGCTCAATAGAGCTCGAGAGATTTGGC 2497
QY 694 GlnSerAlaAlaPheValThrSerThrLeuMetGluGlyGlyValProProSerSer 713
DB 2498 CAGTCGCGAGCTTTGTTGGCTCCACTCTGATGGAGTATGGTGGTGTCTCTCAGTCCGCA 2557
QY 714 SerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLys 733
DB 2558 ACTCGGAGTCTCTTCTGAAAGAACTATCAATGTTATTAAGCTGTGGCTATGAGGACAG 2617
QY 734 ThrAspTrpGlyLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThr 753
DB 2618 ACTGATGGGAACCTGAGATCGGTGGATCTACGGTCTCTGACAGAGACATTTCTCACC 2677
QY 754 GlyPheLysMetHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAla 773
DB 2678 GGATTTCAAGATGCACGCGGAGGCTGGCGGTGCTACTTACTGATGCCCAAGCGGCAGCT 2737
QY 774 PheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAla 793
DB 2738 TTCAAGGGTCTGCCCCCATCAATCTTTCCGACCGTCTTGAACACAGGTGCTCCGGTGGCT 2797
QY 794 LeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsn 813
DB 2798 CTTGGTCCGTGGAGATCCTCTTACGCGGCACATGCCCCCTGTGGTACGGGTAC---GGA 2854
QY 814 GlyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThr 833
DB 2855 GGGCGGCTCAAGTCTCTGGAGAGATTCCGTGATCATCAACACACCATCTACCCGCTCAGC 2914
QY 834 SerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPhe 853
DB 2915 TCCATCCGCTTCTCATCTACTGATCTGCGCGCATCTGTCTGTCTCACCAGAAAGTTC 2974
QY 854 IleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIle 873
DB 2975 ATCATTCAGAGATCAGCAACTTCGCCAGCATCTGTGTTCACTCCCTCTTCATCTCGATC 3034
QY 874 PheAlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArg 893
DB 3035 TTCGCCAGCGCATCTCGAGATGAGTGGAGCGGGTGGGCATCGACAGTGTGGAGG 3094
QY 894 AsnGluGlnPheTrpValIleGlyValSerAlaHisLeuPheAlaValIleGlnGly 913
DB 3095 AACGAGCAGTTCCTGGGTGATCGGGGCATCTCCGCGCACCTCTTCGCGGTGTTCCAGGCG 3154
QY 914 LeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGly 933
DB 3155 CTGCTCAAGGTGTGGCGGCATCGACACCAACTTCACCGTCACTCCAGGCGCTCG--- 3211
QY 934 AspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIlePro 953
DB 3212 GACGAGGACGCGACTTCGCGGAGCTGTACATGTTCAAGTGGACGACGCTCTGTATCCG 3271
QY 954 ProThrThrLeuLeuIleAsnValIleGlyValValAlaGlyIleSerAspAlaIle 973
DB 3272 CCCACCACTCTGATCATCACTGTGTGGCGCTCGTCCGCGCATCTCTACGCCATC 3331
QY 974 AsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpVal 993
DB 3332 AACAGCGATACCACTGTGGGGCGCTCTTCGCGAAGCTCTTCTCGCTCTCTGGGTTC 3391
QY 994 IleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThr 1013
DB 3392 ATCGTCACTGTACCGGTTCTCAAGGGCTCTCAAGGGCTCATGGCAGGACAGACCCCGACC 3451
QY 1014 ValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIle 1033

```

Db 3452 ATCTGCTGCTGGCCATCTCTGCGGCCATCTCTCTTGGTTCGATC 3511
      ::::::::::::::::::::|
Qy 1034 AspProhelleValArgThrLysGlyProAspValargGlnCysGlylleAsnCys 1052
      |||||
Db 3512 GACCCCTTCAACACCGCGTCACTGGCCGGATACCCAGAGCGTGTGCATCAACTGC 3568
      ::::::::::::::::::::|
RESULT 9
US-10-160-719A-57
; Sequence 57, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR FILING DATE: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 3704
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (272)...(3499)
US-10-160-719A-57
Alignment Scores:
Pred. No.: 0 Length: 3704
Score: 3903.50 Matches: 751
Percent Similarity: 78.41% Conservative: 99
Best Local Similarity: 69.28% Mismatches: 151
Query Match: 69.11% Indels: 83
DB: 4 Gaps: 19
US-10-627-132-30 (1-1052) x US-10-160-719A-57 (1-3704)
Qy 32 LeuSerGlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAsp 51
      ::::::::::::::::::::|
Db 308 GTGGCGGCGCAGGTGTGCGCAGATCTGCGCGCAGCGCGTGGCGCACCGCGCGGCGGCGGAC 367
      ::::::::::::::::::::|
Qy 52 LeuPheValAlaCysAsnGluCysGlyPheProValCysArgProCysTyrGluTyrGlu 71
      |||||
Db 368 CTCCTTCAACGCTGCGAGCTGCGCGCTTCCCGGTGCGCGCCCATGCTACGATACGAG 427
      ::::::::::::::::::::|
Qy 72 ArgArgGluGlyThrGlnAsnCysProGlnCysLysThrArgTyrLysArgLeuLysGly 91
      ::::::::::::::::::::|
Db 428 CGCAGGACGCGCACCGCGCGTCCCGCAGTGCAGACTAGTACAGCGGCCACAGGG 487
      ::::::::::::::::::::|
Qy 92 SerProArgValAlaGlyAspAspGluGlu---AspIleAspAspLeuGluHisGlu 110
      ::::::::::::::::::::|
Db 488 AGCCACCACTACACGGTGAGGAAATGAGGATGTGGATGTGTCAGAT----- 535
      ::::::::::::::::::::|
Qy 111 PheAenIleAspAspGluAsnGlnGlnArgGlnLeuGluGlyAsnMet---GlnAsnSer 129
      ::::::::::::::::::::|
Db 536 -----GTGAGTGACTACAACCTAC-----CAAGCATCTGGCAACCGAGATCAGAGCA 583
      ::::::::::::::::::::|
Qy 130 GlnIleThrGluAlaMetLeuHisGlyArgMetSerTyrGlyArgGlyProAspAspGly 149
      ::::::::::::::::::::|
Db 584 AAGATTGCTGAGAGATGCTCACTTGGCGGACAAAC---TCACGCTGGCGAGTATATTGCG 640
      ::::::::::::::::::::|
Qy 150 -----AspGlyAsnAsnThrPro 155
      ::::::::::::::::::::|
Db 641 CTGGCTAAGATATACAGCGCGGTGAAATTTGGGCATGGGAAAGTATGACAGTGGTGCATCCTCCT 700
      ::::::::::::::::::::|
```

```

Qy 156 Gln---IleProProlleIleThrGlySerArgSerValProValSerGlyGluPhePro 174
      ::::::::::::::::::::|
Db 701 CGTGGATATATCCCGTCACCTAACTCATAGCCAG-----ATCTCAGGAGAGATTCTCT 751
      ::::::::::::::::::::|
Qy 175 IleThrAsnGlyTyrGlyHisGlyGluValSerSer-----SerLeuHisLysArg 191
      ::::::::::::::::::::|
Db 752 -----GGAGCTTCCCGCTGATCATATGATGCTCTCTGTTGGGAACATTTGGCAGCGT 802
      ::::::::::::::::::::|
Qy 192 IleHisProTyrProVal-----SerGluProGlySerAlaLysTyrAspGlu 207
      ::::::::::::::::::::|
Db 803 GGACATCAATTTCTTATGTAATCATTTCTCCAAACCCATCGAGGGAGTTCCTCGGTAGC 862
      ::::::::::::::::::::|
Qy 208 LysLysGluValSerTyrLysGluArgMetAspTyrLysSerLys---GlnGlyIle 226
      ::::::::::::::::::::|
Db 863 CTTGGCNAATGTTGCATGGAAAGAGAGGGTGGATGGGAAATGAAGGATAAAGGTGCA 922
      ::::::::::::::::::::|
Qy 227 Leu-----GlyGlyGlyAlaAspProGluAspMet 236
      ::::::::::::::::::::|
Db 923 ATTCCTATGACCAATGGAAACAAGCATTTGCTCCATCAGAGGGCGTGGAGTTGCTGATATT 982
      ::::::::::::::::::::|
Qy 237 AspAla-----AspValAlaLeuAsnAspGluAlaArgGlnPro 249
      ::::::::::::::::::::|
Db 983 GATGCTTCTACTGATTATAACATGGAAGATGCGCTTACTGAATGATGAACCTCGCAACCT 1042
      ::::::::::::::::::::|
Qy 250 LeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleVal 269
      ::::::::::::::::::::|
Db 1043 CTATCTAGAAAGTGGCAATTCCTTCATCCAGATAAATCCGTACAGAAATGGTCATTGTG 1102
      ::::::::::::::::::::|
Qy 270 ValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAsp 289
      ::::::::::::::::::::|
Db 1103 CTACGTTTGGCTGTTCTATGCATATTTCTCGGTACCGTATCACACATCTCTGTGAACAAT 1162
      ::::::::::::::::::::|
Qy 290 AlaIleGlyLeuTyrPheLeuValSerIleCysGluIleTyrPheAlaIleSerTyrIle 309
      ::::::::::::::::::::|
Db 1163 GCATATCCACTGTGGCTTTTATCCGTATATGTGAGATCTGGTTTGTCTTGTCTGGATT 1222
      ::::::::::::::::::::|
Qy 310 LeuAspGlnPheProLysTyrPhePheProIleAspArgGluThrTyrLeuAspArgLeuSer 329
      ::::::::::::::::::::|
Db 1223 TTGGATCAGTTCCCAAGTGGTCCCAANTCAACCGTGAACATACCTTGTAGACTGGCT 1282
      ::::::::::::::::::::|
Qy 330 LeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSer 349
      ::::::::::::::::::::|
Db 1283 TTAAGGTATACCGAGAAGGTGAACCATCTCAATTAGCTCTCTGTGTGATATTTTTGTGAGT 1342
      ::::::::::::::::::::|
Qy 350 ThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeu 369
      ::::::::::::::::::::|
Db 1343 ACTGTGGATCCAAATGAAGGAGCTCTCTTGTCTACTGCAAAATACTGTCTTTCATCTCTT 1402
      ::::::::::::::::::::|
Qy 370 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 389
      ::::::::::::::::::::|
Db 1403 GCTGTCCATATCCGTTGACAGGTATCTTGTCTATGTTTCGGATGATGGAGCTGTATG 1462
      ::::::::::::::::::::|
Qy 390 LeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTyrValProPheCys 409
      ::::::::::::::::::::|
Db 1463 CTGACTTTGATGCTCTCTGAAACTTCAGAGTTTCTAGAAAATGGTTCCGTTCTGT 1522
      ::::::::::::::::::::|
Qy 410 LysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyr 429
      ::::::::::::::::::::|
Db 1523 AAGAAGTACAACTAGAGCTTAGGCGCCGGAATGGTACTTTGTCTCGAAAATTTGATTAC 1582
      ::::::::::::::::::::|
Qy 430 LeuLysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArgGluTyr 449
      ::::::::::::::::::::|
Db 1583 TTGAAGACAAAGTTCAACCTCATTTGTGAAGACCGCGGCGCATGAAGAGAGATAT 1642
      ::::::::::::::::::::|
Qy 450 GluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGlu 469
      ::::::::::::::::::::|
Db 1643 GAAGAAATCAAAAGTTCGTATCAATGCTTTGTAGCAAGGACAAAAGTTTCCGAGAG 1702
      ::::::::::::::::::::|
Qy 470 GlyTyrIleMetLysAspGlyThrProTyrProGlyValAsnAsnThrArgAspHisProGly 489
      ::::::::::::::::::::|
Db 1703 GGATGGATCATGCAAGATGGTATACCTTGGCTGGGAACAATATCTAGGAGCAATCTCTGGA 1762
      ::::::::::::::::::::|
```

QY 490 MetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuPro 509
DB 1763 ATGATTCAGGTTTCTCGGTGTCACAGTGGAGGCTTGACGTTGAAGGCAATGAATTCCT 1822
QY 510 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisIleLysLysAlaGly 529
DB 1823 CGTTTGGTTTATGTCTCTGTAAGAAACGTCCTGGATTCACACATCACAGAGAGCTGGT 1882
QY 530 AlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsn 549
DB 1883 GCCATGAATGCATCTGTCGTATCAGCTGCTTACTAATGGGCAATACATGTTGAAT 1942
QY 550 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeu 569
DB 1943 CTTGATGTGACCATACATCAATATATAGCAAGGCTCTTCGAGAAGCTATGTGCTTCCT 2002
QY 570 MetAspProGlnValGlyArgLysValCysTyrValGlnPhePheGlnArgPheAspGly 589
DB 2003 ATGGACCCAAACCTAGGAAGGATGCTGTATATGTCCTCAATTTCTCAGAGGTTTGAATGGT 2062
QY 590 IleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLys 609
DB 2063 ATTGATAGGAATGACCGATATGCAACACAGGAACACTGTGTTTTCGATATTAACCTGAGA 2122
QY 610 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgGln 629
DB 2123 GGTCTTGACGGCATTCAGAGGGCCAGTTATGTGGAACTGGTGTGTGTTAAACAGAACG 2182
QY 630 AlaLeuTyrGlyTyrAsnProLysGlyProLysArgProLysMetValThrCysAsp 649
DB 2183 GCCTTATATGTTATGAGCCTCCAGTCAAGAAAAAAGCCAGGCTTCTCTCT 2236
QY 650 CysCysProCysPheGlyArgLysLys----- 658
DB 2237 ---TCGCTTTGTGGGGAGGAAAAAGACGCTCAAAATCTAAGAGAGCTCGSAAAGAG 2293
QY 659 -----ArgLysHisAlaLysAspGlyLeuPro-----Glu 668
DB 2294 AAGTCACATAGACACGACAGACTCTGTACCACTATTAAATCTCGAAGATATAGAGAA 2353
QY 669 GlyThrAlaAspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPhe 688
DB 2354 GGGATTGAAGGTTCTCAGTTTGATGATGAGAAATCGCTGATATGTCCTCAATGAGCTTG 2413
QY 689 GluLysArgPheGlyGlnSerAlaPheValThrSerThrLeuMetGluGluGly 708
DB 2414 GAGAAGAGATTGGCCAGTCCAGTGTGTTGTAGCCTCTACTCTCATGGAATATGTTGTT 2473
QY 709 ValProProSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCys 728
DB 2474 GTTCCACAACTGCAACTCCAGAGTCTCTCTGAAAGAGCTATTTCATGTCATCAGCTGT 2533
QY 729 GlyTyrGluAspLysThrAspTyrGlyLeuGluLeuGlyTyrIleTyrGlySerIleThr 748
DB 2534 GGCTATGAGGACAAACTGACTGGGGAAGTGGATGGTGAATCTATGGTTCTGTAC 2593
QY 749 GluAspIleLeuThrGlyPheLysMetHisCysArgGlyTyrArgSerValTyrCysMet 768
DB 2594 GAAGACATCTCACCGGATTCAGATGCATGCTCGAGGCTGGCGATCAATCTACTGCATG 2653
QY 769 ProlLysArgAlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGln 788
DB 2654 CCTAAGCCACAGCTTTTCAAGGGATCTGCTCTATCAACCTTTCGATCGTTTGAATCAA 2713
QY 789 ValLeuArgTyrAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeu 808
DB 2714 GTGCTTCGGTGGCTCTGGTTCCATGAAATCTTTTCAGCAGGCAATGTCCTATGG 2773
QY 809 TyrGlyTyrLysAsnGlnLysThrLeuGluArgPheAlaTyrIleAsnThrThr 828
DB 2774 TATGGCTAT---GGAGGGCGGCTTAAATTCCTGGAGAGATTGCTTATATCAACACACA 2830
QY 829 IleTyrProPheThrSerLeuProLeuAlaTyrCysThrLeuProAlaValCysLeu 848

DB 2831 ATTTATCCACTCACATCAATCCGCTCCTCTGCTACTCATATATCCAGCAGTTTGTCTT 2890
QY 849 LeuThrGlyLysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIleAla 868
DB 2891 CTCACTGGAGAGTTTCATCATCCCAAGATTAGTAACCTAGAGAGTGTGGTTTATATCG 2950
QY 869 LeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValSerIle 888
DB 2951 CTCCTTATCTCAATCTTTGCCACTGGTATCCTTGAGATGAGGTGGTGGTGTGGCATT 3010
QY 889 GluGluTyrTyrArgAsnGluGlnPheTyrValIleGlyGlyValSerAlaHisLeuPhe 908
DB 3011 GATGAATGCTGGAGAACGACGAGCTTCGGGTCAATTTGGTGTATTTCTGCCGATTATTT 3070
QY 909 AlaValValGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThr 928
DB 3071 GCGGTCTTCAGGGTCTCCTGAGGTGCTGCTGTGATTCGACACAGGCTTCACTGTACCC 3130
QY 929 SerLysAlaThrGlyAspGluAspGluPheAlaGluLeuTyrAlaPheLysTyrThr 948
DB 3131 TCTAAGGCCACT---GACGAAGAAGGTGATTTTGGCGAGCTTACATGTTCAAGTGGACA 3187
QY 949 ThrLeuLeuIleProProThrThrLeuLeuIleAsnValIleGlyValValAlaGly 968
DB 3188 ACGTCTTCGATCCCAACCACTATTTTGTATCATCAACCTGGTGGGCTGGTGGCTGGC 3247
QY 969 IleSerAspAlaIleAsnAsnGlyTyrGlnSerTyrGlyProLeuPheGlyLysLeuPhe 988
DB 3248 ATTTCTACGACATCAATAGCGGTACCGTACGATCGGGACCTCTTTTCGGAAGCTCTTC 3307
QY 989 PheAlaPheTyrValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGln 1008
DB 3308 TTTGGTCTCTGGGTGATTTGTCACCTGTACCTTCTCAAGGGCTCATGGGAAGCAG 3367
QY 1009 AsnArgThrProThrValValIleTyrSerIleLeuLeuAlaSerIlePheSerLeu 1028
DB 3368 AACCGCAGCGCGACCATTTGCTGTGGGCTATCTCTTGGCTCGATCTTTTCCCTG 3427
QY 1029 LeuTyrValArgIleAspProPheIleValArgThrLysGlyProAspValArgGlnCys 1048
DB 3428 ATGTGGGTCTGATTCGATCCATTCACACCCGGGCTACTGGCCCTGATATCCGAAATGT 3487
QY 1049 GlyIleAsnCys 1052
DB 3488 GGCATCAACTGC 3499

RESULT 10

US-09-221-013A-9
; Sequence 9, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 3614
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (217)..(3411)

Db 2134 AGTCTTTATCTAAGCTCTGTGGTGA-----TCAGAAAGAAAGAAATTCCAAAGCTAAG 2187
Qy 663 ----- 663
Db 2188 AAAGAGTCGCAAAAGAAATCAGGAGGCATCTGACTCAACTGTTCTCTGATTCAAC 2247
Qy 664 ---AspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAspGlyMetLeu 682
Db 2248 CTCGATGACATAGAGAGGAGTGAAGGTGCTGTTTGTGATGATCAAAAGCGCTCTTA 2307
Qy 683 MetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThrSerThr 702
Db 2308 ATGTCGCAATGAGCTCGAGAGCGATTTGACAGCTGCTGTTTGTGCTCTTCTACC 2367
Qy 703 LeuMetGluGluGlyValProSerSerSerProAlaAlaLeuLeuLysGluAla 722
Db 2368 CTAATGGAATGGTGGTCTCTCTTCAGCAACTCCAGAAACTTTTCAAGAGGCT 2427
Qy 723 IleHisValIleSerCysGlyTyrgluAspLysThrAspTrpGlyLeuLeuGlyTrp 742
Db 2428 ATCCATGTCATTAGTTGGTTATGAGGATAAGTCAGATTGGGAATGAGATTGGATGG 2487
Qy 743 IleTyrglySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArgGlyTrp 762
Db 2488 ATCTATGTTCTGTGACAGAGATATTCGACTGGTTCAAATGATGCCCGCTGGATGG 2547
Qy 763 ArgSerValTyrcysMetProLysArgAlaAlaPheLysGlySerAlaProIleAsnLeu 782
Db 2548 CGATCCATTACTGTCATGCTTAAGCTCCAGCTTTCAGGGTCTGCTCTATCAATCTT 2607
Qy 783 SerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePheSer 802
Db 2608 TCAGATCGTCTGAACCAAGTGTGAGTGGGCTTTAGGTTCAGTTTCAGATTCTCTTCAGT 2667
Qy 803 ArgHisSerProLeuLeuTyrclyTyrglyAsnGlyAsnLeuLysTrpLeuGluArgPhe 822
Db 2668 CGCATTTGCTTATATGTTAGTTTAC---AATGGAGGCTAAATTTCTTGAGAGGTTT 2724
Qy 823 AlaTyrlleAsnThrThreTyrcProPheThrSerLeuProLeuLeuAlaTyrcysThr 842
Db 2725 GCGTATGTGAACACCACCATCTACCTATCCATCTCTCTCTCTCATGTATTGTACA 2784
Qy 843 LeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThrPheAla 862
Db 2785 TTGCTAGCGGTTGTCTCTCCACCAACAGGTTTATTTCCTCAGATTAGTAACATTGCA 2844
Qy 863 SerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArg 882
Db 2845 AGTATGTTGTTCTGTCT 2904
Qy 883 TrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGlnPheTrpValIleGlyGly 902
Db 2905 TGGAGTGGCGTAGGATAGACCAAGTGTGGAGAAACGACGATTTTGGTCAATGGTGA 2964
Qy 903 ValSerAlaHisLeuPheAlaValGlnGlyLeuLeuLysValLeuAlaGlyIleAsp 922
Db 2965 GTATCCGCTCATTTATGCTGTGTTCAAGGTATCTCTCAAGTCTCTCAAGTCTCTGCGTATTGAC 3024
Qy 923 ThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspAspGluPheAlaGluLeu 942
Db 3025 ACAAACTTCACAGTTTACCTCAAAAGCTTCA---GATGAAGACGGAGACTTGTCTGAGTCT 3081
Qy 943 TyrAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeuLeuIleIleAsnVal 962
Db 3082 TACTTGTTCATATGACAACTTCTGATTTCGCCCAACGACGCTCTCATTTGTAACCTTA 3141
Qy 963 IleGlyValValAlaGlySerAspAlaIleAsnAsnGlyTyrglnSerTrpGlyPro 982
Db 3142 GTGGGAGTGTGTCAGGAGTCTCTTATGCTATCAACAGTGGATACCAATCATGGGGACCA 3201
Qy 983 LeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrcProPheLeuLys 1002
Db 3202 CTCCTTGGTAAGTTGTTCTTGGCTTCTGGGTGATTGTTCACTTGTACCTTTCTCTCAAG 3261

RESULT 11

US-10-160-719A-13
; Sequence 13, Application US/10160719A
; Patent No. 6803498

GENERAL INFORMATION:

; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helehtjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13

; LENGTH: 3725
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (179)...(3400)

US-10-160-719A-13

Alignment Scores:

Pred. No.: 0 Length: 3725
Score: 3899.00 Matches: 738
Percent Similarity: 77.19% Conservative: 125
Best Local Similarity: 66.01% Mismatches: 145
Query Match: 69.03% Indels: 110
DB: 4 Gaps: 18

US-10-627-132-30 (1-1052) x US-10-160-719A-13 (1-3725)

Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
Db 179 ATGGCGGCCAACAAAGGGATGGTGGCAGGCTCTCACAAACCCGACAGTTCGTCTATGATC 238
Qy 21 ArgGlyHisGluAsp-----ProLysProLeuArgAlaLeuSerGlyGlnVal 36
Db 239 CGCCACAGCGGACGCGCTGTCCCGCTAAGCCAGAGTGCAGATGGGAGGTC 298
Qy 37 CysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCys 56
Db 299 TGCCAGATTTGTGGCAGACACTGTTGGCGTTTCAGCCACTGGTGTCTTTGTTGCTGTC 358
Qy 57 AsnGluCysGlyPheProValCysArgProCysTyrglyTyrglyGluArgGluGlyThr 76
Db 359 AATGAGTGTGCTTCCCTGTCTGCGCCCTTGTCTATGATGACGACGCAAGAGGGAAC 418
Qy 77 GlnAsnCysProGlnCysLysThrArgTyrglyArgLeuLysGlySerProArgValAla 96
Db 419 CAATGCTGCCCTCAGTGCAGACTAGATACAGAGACAGAAAGGTAGCCCTCGATTCAAT 478
Qy 97 GlyAspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGlu 116

Db 479 GGTGATGATGAGGAGAGATGTTGATGACCTGGCAATGAATTCACACTATAAG----- 532
Qy 117 AsnGlnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeu 136
Db 533 -----CAGGCAAT----- 541
Qy 137 HisGlyArgMetSerTyrGlyArgGlyPro-----AspAspGlyAsp 150
Db 542 -----GGNAGGGCCACAGATGGGAGCTTCAAGAGATGACGCTGAT 583
Qy 151 GlyAsnAsnThrPro-----GlnIleProPheIleThrGlySer 164
Db 584 CTGCTTCATCTGCTCGCATGACCCACACCATCGGATTCACGCTTACAAAGTGGCAA 643
Qy 165 ArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrGlyHisGly----- 182
Db 644 CAG-----ATATCTGGAGAGATCCCTGATGCATCCCTGACCGCTCATCTTATCCGC 694
Qy 183 GluValSerSerSerLeuHisLeuArgIleHisProTyrProValSerGluProGlySer 202
Db 695 AGTCAACATCGAGCTATGTTGATCCAAAGCGTTCAGTTCTCTGTGAGGATTGTG----- 748
Qy 203 AlalysTrpAspGluLysGlyGlu-----ValSerTrpLys 214
Db 749 -----GACCCCTCGAAGGACTTGAATTCCTATGGCTTAATAGTGTGACTGGAAG 799
Qy 215 GluArgMetAspAspTrpLysSerLysGln----- 224
Db 800 GAAAGAGTTGAGAGCTGAGGGGTTAAACAGGACAAAAATATGTTGCAAGTGACTAATAAA 859
Qy 225 -----GlyIleLeuGlyGlyAlaAspProGluAspMetAspAla 238
Db 860 TATCCAGAGGCTAGAGAGACATGAGGGGACTGCTCAATGGAGAGATATGCAA--- 916
Qy 239 AspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSer 258
Db 917 -----ATGGTTGATGATGCACCGCTACCTTGTAGCCGCAATGTCCTCA 967
Qy 259 SerLysValAsnProTyrArgMetValIleValArgLeuValValLeuAlaPhePhe 278
Db 968 AACCAAGCTCAACCTTTACCGGATAGTAATATCTCCGCTTATCATCTGCTGCTTCTC 1027
Qy 279 LeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerIle 298
Db 1028 TTCCAAATATCGATCAGTCATCCAGTCGTAATGCTTATGGAATGTGGCTAGTATCTGTT 1087
Qy 299 IleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPhePro 318
Db 1088 ATCTGTGAGGCTGCTGTTGGCTTGTCTGCTTCTAGATCAGTTCCTCAAAATGGTATCCA 1147
Qy 319 IleAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGluPro 338
Db 1148 ATCAACCGTGACATATCTGCACAGGCTTGCATTGAGGTATGATAGAGGGGAGAGCCA 1207
Qy 339 SerLeuLeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluProPro 358
Db 1208 TCACAGCTGGCTCCCATTTGATGCTTTGTTCAGTACAGTGGATGCCATTGAAGAACCTCCA 1267
Qy 359 LeuValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysVal 378
Db 1268 CTGATCACAGCCAACTGTTTGTTCATCTTGTGTGGATTACCTCTGTGACAAAGTG 1327
Qy 379 SerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThr 398
Db 1328 TCATGCTATGTTCTGATGATGGCTCAGCTATGCTGACTTTTGGCTCTCTCTGAAACT 1387
Qy 399 AlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArgAla 418
Db 1388 GCCGAATTTGCTAGAAAGTGGTTCCTTTTGTGAAGACCAATATTGAACCAAGAGCT 1447
Qy 419 ProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThrPhe 438
Db ----- 458

Db 1448 CCAGAAATTTTACTTTGCTCAAAAAATAGATTACCTGAAGGACAAAAATTCACCTTCATTT 1507
Qy 439 ValGluGluArgArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAla 458
Db 1508 GTTAAGGAAAGACAGCAATGAAGAGAGAGATATGAAGAATTCAAAATTAAGATCAATGCC 1567
Qy 459 LeuValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrPro 478
Db 1568 CTGTTGTTCCAAAGACACAGAAAGTGCCTGAAGAGAGGGTGGACCATGGCTGATGAACCTGCT 1627
Qy 479 TrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSer 498
Db 1628 TGGCTCGGGATTAACCTTAGGACCATCTCTGGCATGATTTCAGGTGTTCTTTGGGGCAGAT 1687
Qy 499 GlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLys 518
Db 1688 GGTGGGCTTGACACTGATGGAATGAATACACAGCTTCTGTCTATGTTCTCTGCTGAAGAAG 1747
Qy 519 ArgProGlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSer 538
Db 1748 AGACCAAGCTTTCAGCATCACAAAGAGGCTGGTGAATGAATGCACCTGATTCTGTGTATCT 1807
Qy 539 AlaValLeuThrAsnAlaPropheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsn 558
Db 1808 GCTGTGCTGACAAATGGTGTCTATCTTCTCAATGTGGATTGTGACCAATTCTTCAATAGC 1867
Qy 559 SerIleValIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysVal 578
Db 1868 AGCAAGCTCTTAGAGAGCAATGTGCTTTCATGATGATCCAGCTCTAGGAGGAAACT 1927
Qy 579 CysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsn 598
Db 1928 TGTTATGTACAAATTCACAAAGATTGATGGCATTGACCTGCACGATCGATATGCTAAT 1987
Qy 599 ArgAsnThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProVal 618
Db 1988 AGGAACATAGTCTCTTTGATATCAACATGAAGGTCTAGATGGCAATTCAGGGTCCAGTC 2047
Qy 619 TyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnPro----- 636
Db 2048 TATGTGGAAACAGATGCTGTTTCAATAGGAGGCTTTGATGATATGATCTCTGTTTGT 2107
Qy 637 ProLysGlyProLysArgProLysMetValThrCysAspCysCysProCysPheGlyArg 656
Db 2108 ACTGAAGCTGATCGGAACCTTAACATTGTTGTTAAAGAGTGTCT-----GGTAGA 2158
Qy 657 LysLysArgLysHisAlaLys----- 663
Db 2159 AGGAAGAGAAAGAACAAAGAGTTATATGATAGTCAAAAGCCGTATTATGAAGAGAACAGAA 2218
Qy 664 -----AspGlyLeuProGluGlyThrAlaAspMetGlyVal 675
Db 2219 TCTTCAGCTCCATCTTTAATCATGAAGACATCGAGAGGGTATTGAA-----GGTTAT 2272
Qy 676 AspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyClnSer 695
Db 2273 GAGGATGAAGGTGCTGCTTATGTCCAGAGGAAATTTGGAGAAACGCTTTGGTCACTCT 2332
Qy 696 AlaAlaPheValThrSerThrLeuMetGluGlyValProProSerSerSerPro 715
Db 2333 CCAATCTTCAATTCATCCACCTTTTATGACTCAAGGTGGCATACCACCTTCAACAAACCA 2392
Qy 716 AlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAsp 735
Db 2393 GCTTCTCTTACTGAAGAGAGCTATCCATGTTATACGCTGTGGGTAGGAGACAAACTGAA 2452
Qy 736 TrpGlyLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPhe 755
Db 2453 TGGGGAAAGAGATTGGCTGGATCTATGTTTCACTTACAGAGGATATTCTGACTGGGTTT 2512
Qy 756 LysMetHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLys 775
Db 2513 AAAATGTCATGAAGAGGCTGGCAATCAATCTACTGTCATGCCACCGACCTTGTTTTCAAG 2572

```
QY 776 GlySerAlaProIleAsnLeuSerArgLeuAenGlnValLeuArgTrpAlaLeuGly 795
Db 2573 GGTCTCTGCACCAATCAATCTTCTGATCGTCTTAATCAGGTCTCGGTGGCTCTTGG 2632
QY 796 SerValGluIlePhePheSerArgHisSerProLeuLeuLeuGlyTrpLysAsnGlyAsn 815
Db 2633 TCAGTGGAAATTCGCTTAGCAGACATGCTCTATATGATGGCTAC---AATGGGCA 2689
QY 816 LeuLysTrpLeuGluArgPheAlaTyrlleAsnThrIleThrProPheThrSerLeu 835
Db 2690 TTGAAGCTTTTGGAGAGGCTGCTTACATTAACACCAATGTTTATCCATCATCTGTT 2749
QY 836 ProLeuLeuAlaTyrcysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMet 855
Db 2750 CCGTATATGCGCTATGTTGCTGCTCTATCTGCTTCTTACCAATAAATTTATCAT 2809
QY 856 ProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAla 875
Db 2810 CCGAGATTAGTAATATGCTGGAAATGTTCTTCTTCTTTTTCCTCCATTTTCGA 2869
QY 876 ThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGlu 895
Db 2870 ACTGCTATATGAGCTCAGATGAGTGTGTGTCATTTGAAGATTGGTGAGAAATGAG 2929
QY 896 GlnPheTrpValIleGlyValSerAlaHisLeuPheAlaValIleGlnGlyLeuLeu 915
Db 2930 CAGTTTGGGTATTGGTGGACCTCTGCCCATCTCTTCCGCGTGTTCAGGCTCTGCTG 2989
QY 916 LysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGlu 935
Db 2990 AAAGTGTGGCTGGGATGATACCAACTTCACAGTTTACCTCAAGGCACT---GATGAG 3046
QY 936 AspAspGluPheAlaGluLeuTyrlaPheLysTrpThrThrLeuLeuIleProProThr 955
Db 3047 GATGCGCATTTGCTGAGCTATATGTTCAAGTGGACGAGTTTCTCATCTCCCTCGACC 3106
QY 956 ThrLeuLeuIleAlaAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsn 975
Db 3107 ACTGTTCTTGTATTAACCTGCTGGAATGTTGGCAGGAAATTCGTATGCCATTAACAGC 3166
QY 976 GlyTyrlleGlnSerTrpGluPheGlyLysLeuPhePheAlaPheTrpValIleVal 995
Db 3167 GCCTACCAATCTCGGGTCCGCTCTTTGGAAAGCTGTCTTCTCATCTGGGTGATCTCTC 3226
QY 996 HisLeuTyrlleProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValVal 1015
Db 3227 CATCTTACCCCTTCTCAAGGCTCTCATGGCAGCAGACCGCACGCCAATCATCTCTC 3286
QY 1016 ValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspPro 1035
Db 3287 ATCGTTTGGTCCATCTCTGCTGCTATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 3346
QY 1036 PheIleValArgThr---LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3347 TTCATCTCCCGACACAGAAAGCTCCCGCTTTGGGGCAATGGTGTGAATCTGC 3400
```

RESULT 12

```
US-10-160-719A-21
; Sequence 21, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjarie, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE OF INVENTION: Therof
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
```

```
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 3725
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (179)....(3400)
US-10-160-719A-21
```

Alignment Scores:

```
Pred. No.: 0 Length: 3725
Score: 3899.00 Matches: 738
Percent Similarity: 77.19% Conservative: 125
Best Local Similarity: 66.01% Mismatches: 145
Query Match: 69.03% Indels: 110
DB: 4 Gaps: 18
```

US-10-627-132-30 (1-1052) x US-10-160-719A-21 (1-3725)

```
QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
Db 179 ATGGCGGCCAACAAAGGGGATGGTGGCAGGCTCTCACAAACCCGACGAGTTCGTATGATC 238
QY 21 ArgGlyHisGluAsp-----ProLysProLeuArgAlaLeuSerGlyGlnVal 36
Db 239 CGCCACGACGCGGCGGCGCTGTCCGCTAAGCCACGAGAGTGCAGATGGGAGGTC 298
QY 37 CysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCys 56
Db 299 TGCCAGATTTCGCGCACACTGTTGGCGTTTCAGCCACTGCTGATGTCTTTGTTGCTGTC 358
QY 57 AsnGluCysGlyPheProValCysArgProCysTyrlleGluTyrlleGluArgGluGlyThr 76
Db 359 AATGAGTGTGCTTCCCTGCTGCGCGCTTGTATGATGATGACGAGCGCAAGGAGGAAC 418
QY 77 GlnAsnCysProGlnCysLysThrArgTyrlleArgLeuLysGlySerProArgValAla 96
Db 419 CAATGCTGCCCTCAGTGCAGACTAGATACAGAGACAGAAAGGTAGCCCTCGAGTTTCA 478
QY 97 GlyAspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGlu 116
Db 479 GGTGATGATGAGGAGGAAGATGTTGATGACCTGACCAATGAATTAATAAG----- 532
QY 117 AsnGlnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeu 136
Db 533 -----CAAGGCAAT----- 541
QY 137 HisGlyArgMetSerTyrlleArgGlyPro-----AspAspGlyAsp 150
Db 542 -----GGGAAGGCCAGAGTGGCAGCTTCAAGGAGATGACGCTGAT 583
QY 151 GlyAsnAsnThrPro-----GlnIleProProIleIleThrGlySer 164
Db 584 CTGTCTTCATCTGCTGCGCATGACCACCATCGGATTCACGCTTACCAAGTGACAA 643
QY 165 ArgSerValProValSerGlyPheProIleThrAsnGlyTyrlleGlyHisGly----- 182
Db 644 CAG-----ATATCTGGAGAGATCCCTGATCATCCCTGACCGGTCATTTCTATCCGC 694
QY 183 GluValSerSerSerLeuHisLysArgIleHisProTyrlleProValSerGluProGlySer 202
Db 695 AGTCCACATCGACCTATGTTGATCCAGCGTTCAGTTCCTCTGTGAGGATTTGTG----- 748
QY 203 AlaLysTrpAspGluLysLysGlu-----ValSerTrpLys 214
Db 749 -----GACCCCTCGAAGGACTTGAATTCCTTATGGCTTAAATAGTGTGACTGGAAG 799
QY 215 GluArgMetAspTrpLysSerLysGln----- 224
Db 215 ----- 224
```


800 GAAAGAGTTGAGAGCTGGAGGGTTAAACAGACGACAAAAATATGTTGCAAGTCACTAATAAA 859
225 -----GlyIleLeuGlyGlyAlaAspProGluAspMetAspAla 238
860 TATCCAGAGGCTAGAGAGACATGGAGGGGACTGGCTCAAAATGGAGAAGATATGCAA--- 916
239 AspValAlaLeuAenAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSer 258
917 -----ATGGTTGATGATGACAGCTACCTTTGAGCGCATTTGGCCAAATTCCTCA 967
259 SerLysValAenProTyrArgMetValIleValValArgLeuValValLeuAlaPhePhe 278
968 AACCAAGCTCAACCTTTACCGGATAGTAATCAATCTCCGCTCTTATCATCTGCTGCTTCTTC 1027
279 LeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerIle 298
1028 TTCCAATATCGTATCAGTCATCCAGTCGTAATGCTTATGGATTGTGGCTAGTATCTGTT 1087
299 IleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPhePro 318
1088 ATCTGTGAGGCTCGTTTGGCTTGTCTGGCTTCTAGATCAGTCCCAAAATGATATCCA 1147
319 IleAspArgGluTrpTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGluPro 338
1148 ATCAACGCTGAGACATATCTCGACAGGCTTGCATTGAGGTATGATAGAGAGGAGACCA 1207
339 SerLeuLeuSerAlaValAspLeuPheValSerTrpValAspProLeuLysGluProPro 358
1208 TCACAGCTGGCTCCCATTTGATGCTTTGTCTAGTACAGTGGATCCATTGAAGGAACCTCCA 1267
359 LeuValThrAlaAenThrValLeuSerIleLeuAlaValAspTyrProValAspLysVal 378
1268 CTGATCAGACGACACACTGTTTGTCCATCTTGTCTGGATTACCTGTTGACAAAGTG 1327
379 SerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThr 398
1328 TCATGCTATGTTCTGATGATGGCTCAGCTATGCTGCTTTTGTGATGCTCTCTCTGAAACT 1387
399 AlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArgAla 418
1388 GCCGAATTTGCTAGAAATGGGTTCCCTTTTGTGAAGAAGCAATATTGAACCAAGAGCT 1447
419 ProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThrPhe 438
1448 CCAGNATTTTACTTTGCTCAAAAATAGATTACCTGAAGGACMAAAATCAACCTTCATTT 1507
439 ValGlnGluArgArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAla 458
1508 GTTAAGGAAAGACGAGCAATGAAGAGAGATGAAGAATTCMAAAATGAAGATCAATGCC 1567
459 LeuValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrPro 478
1568 CTTGTTGCCAAGACACAGAAATGGCTGAAGAGGGGTGGACATGGCTGATGGAACTGCT 1627
479 TrpProGlyAsnAenThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSer 498
1628 TGGCCTGGGAATTAACCTTAGGACCATCTCGCATGATTACAGTGTCTTGGGACACAGT 1687
499 GlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLys 518
1698 GGTGGGCTTGACACTGGAAATGAATTAACCACTCTTGTCTATGCTCTCTCGTGAAGAAG 1747
519 ArgProGlyPheGlnHisLysLysValAlaGlyAlaMetAsnAlaLeuIleArgValSer 538
1748 AGACAGAGGCTTTCAGCATCACAGAAGGCTGGTGAATGAATGAATGCACTGATTCGTATCT 1807
539 AlaValLeuThrAsnAlaProPheMetLeuAenLeuAspCysAspHisTyrIleAsnAsn 558
1808 GCTGTGCTGACAAATGGTGCCTATCTTCTCAATGTGGATTGTGACCAATTACTTCAATAGC 1867
559 SerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysVal 578
1868 AGCAAGAGCTCTTAGAGAAGCAATGTGCTTTCATGATGGATCCAGCTCTAGGAAGGAAAACT 1927

Qy 579 CysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsn 598
Db 1928 TGTATTGTACATTTCCACAAAGATTTGATGGCATTTGATGGCATTCGATATGCTAAT 1987
Qy 599 ArgAsnThrValPhePheAspIleAenMetLysGlyLeuAspGlyIleGlnGlyProVal 618
Db 1988 AGGAACATAGTCTTCTTTGATATCAACATGAAGGCTAGATGGCATTCAGGCTCCAGTC 2047
Qy 619 TyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnPro----- 636
Db 2048 TATGTGGGAACAGAGTGTCTTCAATAGCAGGCTTTGATGGATGATATCTCTGTTTG 2107
Qy 637 ProLysGlyProLysArgProLysMetValThrCysAspCysCysProCysPheGlyArg 656
Db 2108 ACTGAAGCTATCTGGAACCTAACATTTGTTGTAAGAGCTGCTGT-----GGTAGA 2158
Qy 657 LysLysArgLysHisAlaLys----- 663
Db 2159 AGGAAGAGAAAGAACCAAGAGTTATATGGATAGTCAAAAGCCGTATTATGAAGAGACAGAA 2218
Qy 664 -----AspGlyLeuProGluGlyThrAlaAspMetGlyVal 675
Db 2219 TCTTACGCTCCCATCTTTAACATGGAGACATCGAGAGGGTATTGAA-----GGTTAT 2272
Qy 676 AspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSer 695
Db 2273 GAGGATGAAGGTCAGTGTCTTATGTCCAGAGGAAATTTGAGAAACGCTTTGGTCAGTCT 2332
Qy 696 AlaAlaPheValThrSerThrLeuMetGluGluGlyValProProSerSerSerPro 715
Db 2333 CCAATCTTCACTGCATCCACCTTTATGACTCAAGGTGGCATACCACTTCAACAAACCCA 2392
Qy 716 AlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAsp 735
Db 2393 GCTTCTCTACTGAAGGAGCTATCCATGTTATCAGCTGTGGTACGAGGACAAACCTGAA 2452
Qy 736 TrpGlyLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPhe 755
Db 2453 TGGGAAAGAGATTTGCTGATCTATGTTTACAGGATATTTCTGACTGGGTTT 2512
Qy 756 LysMetHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLys 775
Db 2513 AAAATGCATCAAGAGGCTGGCAATCACTACTGCATGCCACCAACCTTTGTTCAAG 2572
Qy 776 GlySerAlaProIleAenLeuSerAspArgLeuAsnGlnValLeuArgTyrAlaLeuGly 795
Db 2573 GGTTCCTGCACCAATCAATCTTCTGATCGTCTTATCAGGTGCTCCGTTGGGCTCTTGGG 2632
Qy 796 SerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsn 815
Db 2633 TCAGTGGAAATTCGCTTAGCAGACATGTTCTATATGTTATGGTATGCTAC---AATGGGCGA 2699
Qy 816 LeuLysTrpLeuGluArgPheAlaTyrIleAenThrThrIleTyrProPheThrSerLeu 835
Db 2690 TTGAAGCTTTTGGAGAGGCTGGCTTACATTAACACCATTTGTTATCCAATCACAATCTGTT 2749
Qy 836 ProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMet 855
Db 2750 CCGCTTATCGCTATTTGTGCTTCTCGCTATCTCTTCTTACCAATAAATTTATCATTT 2809
Qy 856 ProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAla 875
Db 2810 CCTGAGATTAGTAATTATGCTGGAAATGTTCTTCTTCTTTTTCCTCCATTTTCGCA 2869
Qy 876 ThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGlu 895
Db 2870 ACTGATATTTGGAGCTCAGATGGAGTGGTGTGGCATTTGAAGATTGGTGAGAAATGAG 2929
Qy 896 GlnPheTrpValIleGlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeu 915
Db 2930 CAGTTTTGGGTTATTTGGGCACCTCTGCTCCCATCTCTTCGGGGTGTTCACGGGCTCTGCTG 2989

QY 916 LysValLeuAlaGlyIleAspThrPheThrValThrSerIleAlaThrGlyAspGlu 935
Db |||||
2990 AARGTGTGGCTGGGATTCATACCACTTCACAGTTACCTCAAGGCACTCT---GATGAG 3046
QY 936 AspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThr 955
Db |||||
3047 GATGGCGACTTGGCTGAGCTATATGTTCAGTGGACAGTTGCTCATCTCCGACC 3106
QY 956 ThrLeuLeuIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsn 975
Db |||||
3107 ACTGTTCTTGTCTATTAACCTGGTGGATGGTGGCAGGAATTCGTATGCCATTAAACAGC 3166
QY 976 GlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleVal 995
Db |||||
3167 GCTACCAACTCTGGGTCGGCTCTTTGGAAGCTGTTCTTCGATCTGGGTGATCTCTC 3226
QY 996 HisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValVal 1015
Db |||||
3227 CATCTCTACCCCTTCTCAAGGCTCTCATGGGCGAGCAGAACCGCACCAATCTGTC 3286
QY 1016 ValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspPro 1035
Db |||||
3287 ATCGTTTGGTTCATCTCTCGTCTATCTTCTCTGCTGCTGGGTGAAGATCGATCT 3346
QY 1036 PheIleValArgThr---LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db |||||
3347 TTCATCTCCCGACACAGAAAGCTGCGGCTTGGGGCAATGGTGTGAAGTGC 3400

RESULT 13

US-10-160-719A-41
; Sequence 41, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Drugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: Thereof
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 3725
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (179)...(3400)
US-10-160-719A-41

Alignment Scores:
Pred. No.: 0 Length: 3725
Score: 3899.00 Matches: 738
Percent Similarity: 77.19% Conservative: 125
Best Local Similarity: 66.01% Mismatches: 145
Query Match: 69.03% Indels: 110
DB: 4 Gaps: 18

US-10-627-132-30 (1-1052) x US-10-160-719A-41 (1-3725)

QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
Db |||||
179 ATGGCGGCCAACACAGGGATGGTGGCAGCTCTCAACACCGCAACGAGTTCGTATGATC 238
QY 21 ArgGlyHisGluAsp-----ProLysProLeuArgAlaLeuSerGlyGlnVal 36
Db |||||

Db 239 GCCACGACGGCGACGCCCTGTCCGCTAAGCCACGAAGAGTGCATGGCGCAGGTC 298
QY 37 CysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCys 56
Db |||||
299 TGCAGATTTGGCGCACACTGTTGGCGTTTCAGCCACCTGGTGATGCTTTGTTGCTTGC 358
QY 57 AsnGluCysGlyPheProValCysArgProCysTyrGluTyrGluArgGluGlyThr 76
Db |||||
359 AATGAGTGTGCTTCCCTGTCTGCGCCCTTGTCTATGATGATGACGCGCAAGGAAGAAC 418
QY 77 GlnAsnCysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAla 96
Db |||||
419 CAATCTCCCTCAGTGCACACTAGATACAGAGACAGAAAGGTAGCCCTCGAGTTCA 478
QY 97 GlyAspAspAspGluAspIleAspLeuGluHisGluPheAsnIleAspAspGlu 116
Db |||||
479 GGTGATGATGAGGAGGAAGATGTTGATGACCTGGACAATGAATTCACACTATAAG---- 532
QY 117 AsnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeu 136
Db |||||
533 -----CAAGGCAAT----- 541
QY 137 HisGlyArgMetSerTyrGlyArgGlyPro-----AspAspGlyAsp 150
Db |||||
542 -----GGGAAGGGCCAGAGTGGCAGCTTCAAGGAGATGACCTGAT 583
QY 151 GlyAsnAsnThrPro-----GlnIleProProIleIleThrGlySer 164
Db |||||
584 CTGCTTTCATCTGCTCGCCATGACCACACCATCGGATTCACAGCTTACAAGTGGACAA 643
QY 165 ArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrGlyHisGly----- 182
Db |||||
644 CAG-----ATATCTGGAGAGATCCCTGATGATCCCTCGACCGCATTTCTATCCGC 694
QY 183 GluValSerSerLeuHisLysArgIleHisProTyrProValSerGluProGlySer 202
Db |||||
695 AGTCCACATCGAGCTATGTTGATCCAGCGCTTCCAGTTCCTGTGAGAGTTGTG----- 748
QY 203 AlaLysTrpAspGluLysLysGlu-----ValSerTrpLys 214
Db |||||
749 -----GACCCCTCGAAGGACTTGAATTCCTATGGGCTTAATAGTGTGTTGACTGGAAG 799
QY 215 GluArgMetAspAspTrpLysSerLysGln----- 224
Db |||||
800 GAAAGAGTTGAGAGCTGGAGGGTTAAACAGGACAAAAATATGTTCCAAAGTGAATAAAA 859
QY 225 -----GlyIleLeuGlyGlyAlaAspProGluAspMetAspAla 238
Db |||||
860 TATCAGAGGCTAGAGGACATGGAGGGGACTGGCTCAATGGAGAGATATGCA----- 916
QY 239 AspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSer 258
Db |||||
917 -----ATGGTTGATGATGACCGCTACCTTTGAGCCGCAATGTGCCAATTTCTTCA 967
QY 259 SerLysValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPhePhe 278
Db |||||
968 AACCAAGCTCAACCTTTACCGGATAGTAATCATCTCCGCTTATCATCTCTGCTGCTTCT 1027
QY 279 LeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerIle 298
Db |||||
1028 TTCCATATATCATCATGATCCAGTGGTATGCTTATGATGATGCTGCTGCTGCTGCTGCT 1087
QY 299 IleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPhePro 318
Db |||||
1088 ATCTGTGAGGCTGTGTTGCTTCTGCTGGCTTCTAGATCAGTTCCTCCCAAAATGGTATCCA 1147
QY 319 IleAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGlyGluPro 338
Db |||||
1148 ATCAACCGTGAGACATATCTCAGAGGCTTGCATGAGGTATGATAGAGAGGAGAGCA 1207
QY 339 SerLeuLeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluProPro 358
Db |||||
1208 TCACAGCTGGCTCCCATTTGATGTCTTTGTGCTAGTACAGTGGATCCATTAAGGAACCTCCA 1267

Qy 359 LeuValThrAlaAanThrValLeuSerIleLeuAlaValAspTyrProValAspLysVal 378
Db 1268 CTGATCAGCCACACGCTGTTTGTGCACTTCTGCTGGATACCCCTGTTGACAAAGTG 1327
Qy 379 SerCysTyrValSerAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThr 398
Db 1328 TCATGCTATGTTCTGATGATGGCTCAGCTATGCTGCTTTGAGTCTCTCTCGAAACT 1387
Qy 399 AlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArgAla 418
Db 1388 GCCGAATTGCTGAAAGTGGGTTCCCTTTTGTGAAGACACAATATTGAACCAAGAGCT 1447
Qy 419 ProGluPheTyrPheSerLeuLysValAspTyrIleLysAspLysValGlnProThrPhe 438
Db 1448 CCAGAAATTTACTTTGCTCAAAAATAGATTACCTGGAAGGACAAAATTCACACCTTCATTT 1507
Qy 439 ValGlnGluArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAla 458
Db 1508 GTTAAGGNAAGACGAGCAATGAAGAGAGATATGAGNAATTCAAAATAAGNAATCAATGCC 1567
Qy 459 LeuValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrPro 478
Db 1568 CTTGTGTCAAAGCACAGAAAGTGCCTGAAGAGGGGTGGACCATGGCTGATGGAACTGCT 1627
Qy 479 TrpProGlyAanAanThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSer 498
Db 1628 TGGCTGGGAATAACCCCTAGGAGCAATCCTGGCATGATTCAGGTGTTCTTGGGGCACAGT 1687
Qy 499 GlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLys 518
Db 1688 GGTGGGCTTGACATGATGGAATGAATTACACGCTCTGTCTATGCTCTCTCGTGAAGAG 1747
Qy 519 ArgProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSer 538
Db 1748 AGACAGGCTTTCAGCATCACAAAGAGCTGGTGCATGAATGAATGCATGATTCGTGTATCT 1807
Qy 539 AlaValLeuThrAsnAlaProPheMetLeuAanLeuAspCysAspHisTyrIleAsnAsn 558
Db 1808 GCTGTGCTGACAAATGGTGCCATCTCTCAATGGGATGTGACCATTAATCTCAATAGC 1867
Qy 559 SerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysVal 578
Db 1868 AGCAAGCTCTTAGAAGACATGTGCTTCATGATGATCCAGCTCTAGGAAGGAAACT 1927
Qy 579 CysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsn 598
Db 1928 TGTATGTACAAATTTCCACAAAGATTTGATGGCATTTGATGGCATCGATATGCTAAT 1987
Qy 599 ArgAanThrValPheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProVal 618
Db 1988 AGGAACATAGTCTCTTTGATATCAACATGAAGGCTCTAGATGGCAATTCAGGTTCCAGTC 2047
Qy 619 TyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnPro 636
Db 2048 TATGTGGGACAGAGTGTGTTCAATAGGCAGGCTTTGTATGGATATGATCCCTGTTTG 2107
Qy 637 ProLysGlyProLysArgProLysMetValThrCysAspCysCysProCysPheGlyArg 656
Db 2108 ACTGAAGCTGATCTGGAACCTAACATTTGTTTGAAGAGCTGCTGT-----GSTAGA 2158
Qy 657 LysLysArgLysHisAlaLys----- 663
Db 2159 AGGAAGAGAAAGAACCAAGAGTTATATGGAATAGTCAAGCCGTATTATGAAGAGAACAGAA 2218
Qy 664 -----AspGlyLeuProGluGlyThrAlaAspMetGlyVal 675
Db 2219 TCTTCAGCTCCCATCTTTTAACATGGAACATCCGAGGAGGTTATTTGA-----GTTAT 2272
Qy 676 AspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSer 695
Db 2273 GAGGATGAAGGTCAGTGTCTATGTGCCAGAGAAATTTGGAGAAACGCTTTTGGTCAGTCT 2332

Qy 696 AlaAlaPheValThrSerThrLeuMetGluGluGlyValProProSerSerSerPro 715
Db 2333 CCATCTTTCATTCATCCACCTTATATGACTCAAGTGGCATACCACTTCAACAACCCCA 2392
Qy 716 AlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAsp 735
Db 2393 GCTTCTCTACTGAAGGAAGCTATCCATCTTATCAGCTGTGGGTACGAGGACAAACTGAA 2452
Qy 736 TrpGlyLeuLeuLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPhe 755
Db 2453 TGGGAAAAAGAGATTTGGCTGATCTATGGTTTCAGTTACAGAGGATATCTGACTGGGTTT 2512
Qy 756 LysMetHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLys 775
Db 2513 AANAATGCATGCAAGAGCTGGCAATCAATCTACTGATGCCACCAACGACCTGTTTCAAG 2572
Qy 776 GlySerAlaProIleAanLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGly 795
Db 2573 GGTTCGTGCACCAATCAATCTTCTGATCGCTTAATCAGGTGCTCGGTGGGCTCTTGGG 2632
Qy 796 SerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsn 815
Db 2633 TCAGTGGAAATTCGCTTAGCAGACATGTCTTATATGATGGCTAC-----AATGGCGCA 2689
Qy 816 LeuLysTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeu 835
Db 2690 TTGAAGCTTTTGGAGAGGCTGGCTTACATTAACACCAATGTTTATCCAAATCACATCTGT 2749
Qy 836 ProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMet 855
Db 2750 CCGCTTATCGCTATTTGTGTGCTTCTGCTATCTGCTTCTTACCAATAATTTATCATT 2809
Qy 856 ProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAla 875
Db 2810 CCTGAGATTAGTAATATGCTGGAATGTTCTTCAATCTCTTTTGTGCTCAATTTTCGCA 2869
Qy 876 ThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGlu 895
Db 2870 ACTGTTATATGGAGCTCAGATGGAGTGGTGGTGGCATTGAAGATTGGTGAGAAATGAG 2929
Qy 896 GlnPheTrpValIleGlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeu 915
Db 2930 CAGTTTTGGGTTATTTGGTGCCATCTCTGCGCATCTCTTCGCGGTGTTCCAGGGTCTGCTG 2989
Qy 916 LysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGlu 935
Db 2990 AAGTGTGGCTGGATTGATACCACTTCAGATTACCTCAAGGATCT---GATGAG 3046
Qy 936 AspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThr 955
Db 3047 GATGGGACITTTGCTGAGCTATATGTGTTCAAGTGGACCAAGTTGCTCATCCCTCCGACC 3106
Qy 956 ThrLeuLeuIleIleAanValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsn 975
Db 3107 ACTGTTCTGTCAATTAACCTGGTGGAAATGGTGCGAGAAATTCGTATGCCATTAACAGC 3166
Qy 976 GlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleVal 995
Db 3167 GGCTACCACATCCTGGGGTCCGCTCTTTGGAAAGCTGTTCTTCTCGATCGGTGATCCTC 3226
Qy 996 HisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAanArgThrProThrValVal 1015
Db 3227 CATCTTACCCCTTCTCAGGGTCTCATGGGAGGAGCAACCCGACGCCAACCAATCTGC 3286
Qy 1016 ValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspPro 1035
Db 3287 ATCGTTTGGTCCATCTCTTGGCTCTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3346
Qy 1036 PheIleValArgThr---LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3347 TTCATCTCCCGACACAGAAAGCTGCCCTTGGGGCAGATGTTGGTGTGAATCTGC 3400

RESULT 14

498	Qy	SerGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlu	517
1693	Db	AGTGGTGGGCTCGACACTGATGAAATGAGTTACCAACGCTCTTGCTCTATGCTCTCGTGAA	1752
518	Qy	LysArgProGlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgVal	537
1753	Db	AAGAGACCAGGCTTTCAGCATCAACAAGAGCGTGGTGCATGAATGCCTGATTCGTGTA	1812
538	Qy	SerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsn	557
1813	Db	TC TGCTGTGCTGACAAATGGTGGCTATCTCTCAATGTGGATGGACCACTTACTTCAAT	1872
558	Qy	AsnSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLys	577
1873	Db	AGCAGCAAAAGCTCTTAGAGAAGCAATGTGCTTCATGATGGATCCGGCTCTAGGAAGAAA	1932
578	Qy	ValCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAla	597
1933	Db	ACTTGTATGTACAATTTCCACAGAGATTGTGATTCGCAATTGACTTGCACGATCGATGCT	1999
598	Qy	AsnArgAsnThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyPro	617
1993	Db	AATCGGAACATAGTTTCTTTTGATATCAACATGAAGGCTCGATGCGCAATTCAGGTCCA	2052
618	Qy	ValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnPro---	636
2053	Db	GTTTACGTGGACAGGATGCTGTTTCAATAGACAGGCTTTGTATGGATACGATCTGTT	2112
637	Qy	---ProLysGlyProLysArgProLysMetValThrCysAspCysCysProCysPheGly	655
2113	Db	TTGACTGAAGCTGATCTCGAGCCAAACATGTTTATTAAAGAGCTGCTGT-----GGT	2163
656	Qy	ArgLysLysArgLysHisAlaLys-----	663
2164	Db	AGAGGAGAAAAGAAACAGAGTTATATGGATAGTCAAGACCGGTATATGAAGAGAAC	2223
664	Qy	-----AspGlyLeuProGluGlyThrAlaAspMetGly	674
2224	Db	GAATCTTCAGCTCCCATCTTCAATATGGAAGACATCGAAGAGGATATTGAA-----GGT	2277
675	Qy	ValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGlnLysArgPheGlyGln	694
2278	Db	TACGAGGATGAAGAGCTCAGTCTGCTTATGCTCCAGAGGAAATTTGGAGAAACGCTTTGGTCAG	2337
695	Qy	SerAlaAlaPheValThrSerThrLeuMetGluGluGlyValProProSerSerSer	714
2338	Db	TCTCCTATTTTCATGTCATCCCTTTATGACACAGGTGGCATACCACTTCAACAAC	2397
715	Qy	ProAlaAlaLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThr	734
2398	Db	CCAGCTTCTTACTTAAGGAAGCTATCCATGTCATCATGTTGGATATGAGGACAAAAC	2457
735	Qy	AspTrpGlyLeuGluLeuGlyTyrPileTyrGlySerIleThrGluAspIleLeuThrGly	754
2458	Db	GAATGGGGAAAAAGAGATTGGCTGGATCTATGTTTTCAGTAACGGAGGATATTTCACTGGG	2517
755	Qy	PheLysMetHisCysArgGlyTyrArgSerValTyrCysMetProLysArgAlaAlaPhe	774
2518	Db	TTTAAATATGCATGCAAGGGGCTGGCAATCAATCTACTGTCATGCCACCAACGACCTTGTTTC	2577
775	Qy	LysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTyrAlaLeu	794
2578	Db	AAGGGTTCGACCAACAATCTTTCCGATGCTCTTAATCAGGTGCTCCGTTGGGCTCTT	2637
795	Qy	GlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGly	814
2638	Db	GGGTCAGTGGAAATCTGCTTTAGTAGACATTTGTGCTCTCTCTGATGGTTAC---AATGGA	2694
815	Qy	AsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSer	834
2695	Db	CGATTGAAGCTTTTGGAGAGGCTGGCTTACATCAACACTATGTATATATCCATCACTCC	2754
835	Qy	LeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIle	854

Db	2755	ATTCCGCTTATTGCCATTATGTGTGTGCTTCCGCTATCTGCTCTCTTACCAATAAATTATC	2814
Qy	855	MetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePhe	874
Db	2815	ATTCTCTGAGATTAGCAATATATGCTGGAGATCTCTCTTCTTCTTTTGGCTCCATTTT	2874
Qy	875	AlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsn	894
Db	2875	GCCACTGGTATATTGGAGCTTAGATGAGTGGTGTGGCATTCGAAGATTGGTGGAGAAAT	2934
Qy	895	GluGlnPheTrpValIleGlyValSerAlaHisLeuPheAlaValGlnGlyLeu	914
Db	2935	GAGCAGTTTGGGTATTTGGTGGCACTCTGCCCATCTCTTCCAGTGTTCAGGGTCTG	2994
Qy	915	LeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAsp	934
Db	2995	CTGAAGTGTGGCTGGGATTGATACCACTTCACAGTTACCTCAAAGGCATCT--GAT	3051
Qy	935	GluAspAspGluPheAlaGluLeuTyrrAlaPheGlyTrpThrThrLeuLeuIleProPro	954
Db	3052	GAGGATGGCAGCTTTGCTGAGCTATATGTCTCAAGTGGACCAAGTTTGCCTCAITTC	3111
Qy	955	ThrThrLeuLeuIleLeuValIleGlyValValAlaGlyIleSerAspAlaIleAsn	974
Db	3112	ACCAGTCTTGTGCATTAACCTGGTGGGAATGGTGGCAGGAATTTCTTATGCCATTAAC	3171
Qy	975	AsnGlyTyrrGlnSerTrpGlyProLeuPheGlyLysLeuPheAlaPheTrpValIle	994
Db	3172	AGTGGCTACCAATCTGGGTCTGCTCTTTGGAAAGCTGTCTCTCGATCTGGTGATC	3231
Qy	995	ValHisLeuTyrrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrVal	1014
Db	3232	CTCCATCTCTACCGCTTCTCAAGGGTCTCATGGGAAGGAGAACCGCACCAACAATC	3291
Qy	1015	ValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAsp	1034
Db	3292	GTCATTCTGGTCCATCTCTTGTGCATCTATCTCTCTGTGTGGTGAAGATCGAT	3351
Qy	1035	ProPheIleValArgThr---LysGlyProAspValArgGlnCysGlyIleAsnCys	1052
Db	3352	CCTTTCATCTCCCGACACAGAAAGTGTGCTTGGGGCAATGTGGCGTCAACTGC	3408
RESULT 15			
US-10-160-719A-53			
; Sequence 53, Application US/10160719A			
; Patent No. 6803498			
; GENERAL INFORMATION:			
; APPLICANT: Dhugga, Kanwarpal S.			
; APPLICANT: Helentjaris, Timothy G.			
; APPLICANT: Bowen, Benjamin A.			
; APPLICANT: Wang, Xun			
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses			
; TITLE OF INVENTION: Thereof			
; FILE REFERENCE: 0864C			
; CURRENT APPLICATION NUMBER: US/10/160,719A			
; PRIOR FILING DATE: 2002-06-03			
; PRIOR APPLICATION NUMBER: US 60/096,822			
; PRIOR FILING DATE: 1998-08-17			
; PRIOR APPLICATION NUMBER: US 09/371,383			
; PRIOR FILING DATE: 1999-08-06			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 53			
; LENGTH: 3753			
; TYPE: DNA			
; ORGANISM: Zea mays			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (184)...(3408)			
US-10-160-719A-53			
Alignment Scores:			

Pred. No.: 0 Length: 3753
Score: 3895.50 Matches: 735
Percent Similarity: 76.94% Conservative: 126
Best Local Similarity: 65.68% Mismatches: 147
Query Match: 68.97% Indels: 111
DB: 4 Gaps: 17

US-10-627-132-30 (1-1052) x US-10-160-719A-53 (1-3753)

QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGlnLeuValLeu 20
DB 184 ATGGCGGCCAACAAAGGGATGTTGGCGGCTCGCAACCGCAACGAGTTGCTCATGATC 243
QY 21 ArgGlyHisGluAsp-----ProLysProLeuArgAlaLeuSerGlyGlnVal 36
DB 244 CCCCACGACGCGCGATGTGCGCGGCTCGGCTAAGCCACAAAGAGTGCAGATGACAGGTC 303
QY 37 CysGluLeuLeuValGlyLeuThrValAspGlyAspLeuPheValAlaCys 56
DB 304 TCCAGATTGCGGTGACTCTGTGGGTCTCAGCCACTGGGTGATGCTTTGTCCTGC 363
QY 57 AsnGluCysGlyPheProValCysArgProCysTyrGluTyrGluArgGluGlyThr 76
DB 364 AATGAGTGTGCTTCCCTGTCTGCGCCCATGCTATGATGATGAGCGCAAGGAGGGAAC 423
QY 77 GlnAsnCysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAla 96
DB 424 CAATGCTGCCCCAGTGCAGACTAGATACAAAGACAGCAAGAGGTAGCCCTCGAGTTCA 483
QY 97 GlyAspAspGluGluAspLeuAspLeuGluHisGluPheAsnLeuAspGlu 116
DB 484 GGTGATGAGGATGAGAGATGATGATGACCTAGACAAATGAATCAACTAC----- 534
QY 117 AsnGlnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnLeuThrGluAlaMetLeu 136
DB 534 ----- 534
QY 137 HisGlyArgMetSerTyrGlyArgGlyPro-----AspAspGlyAsp 150
DB 535 -----AAGCAAGGAGTGGGAAGGCCAGAGTGGCAACTGCAAGGAGATGATGCTGAT 588
QY 151 GlyAsnAsnThrPro-----GlnLeuProIleLeuThrGlySer 164
DB 589 CTGCTTTCATCTGCTGCCCATGAGCCACATCATCGGATTCACGCTGACAGCGGTCAA 648
QY 165 ArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrGlyHisGly----- 182
DB 649 CAG-----ATATCTGGAGAGATCTCTGATGCTTCCCTGACCGTCAATCTATCCGC 699
QY 183 GluValSerSerSerLeuHisLysArgLeHisProTyrProValSerGluProGlySer 202
DB 700 AGTCCAACATCAGCATGATGTTGATCCAAGCGTCCCAAGTCTCTGAGGATTGTG----- 753
QY 203 AlaLysTrpAspGluLysGlu-----ValSerTrpLys 214
DB 754 -----GACCCCTCGAAGGACTTGAATTCCTATGGGCTTAATGATGTTGACTGGAAG 804
QY 215 GluArgMetAspAspTrpLysSerLysGln----- 224
DB 805 GAAAGATTGAGAGCTGGAGGTTAAACAGGCAAAATATGATCAAGTGAATAATAA 864
QY 225 -----GlyLeuGlyGlyGlyAlaAspProGluAspMetAsp 237
DB 865 TATCCAGAGGCTAGAGGAGGACATGAGGCGGACTGGCTCAAAATGGAGAGATATGCAA 924
QY 238 AlaAspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAla 257
DB 925 -----ATGGTTGATGATCAGCGCTACCTTTGAGCCGCTATCGTGCCTATCC 972
QY 258 SerSerLysValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPhe 277
DB 973 TCAACACAGCTCAACTTTACCGGGTAGTGATCATCTCCGCTTATCATCTCTGCTTC 1032

QY 278 PheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSer 297
DB 1033 TTCTTCCAGTATCGTGTGCTAGTCATCCAGTCCGCTGATGCTTATGGATTATGGCTAGTATCT 1092
QY 298 IleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPhe 317
DB 1093 GTTATCTCGAGGCTCGTTTGGCTTCTTGGCTTCTAGATCAGTTCCCAANATGGTAT 1152
QY 318 ProIleAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGlu 337
DB 1153 CCAATCAACCGTGAGACATATCTTGACAGGCTTGCTAGGATGATAGAGAGGAGAG 1212
QY 338 ProSerLeuLeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluPro 357
DB 1213 CCATCACAGCTGGCTCCCATGATGCTCTCGTACGTACAGTGGATCCATTTGAAGAACCT 1272
QY 358 ProLeuValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLys 377
DB 1273 CCACTGATCAGCCACACACTGTTTGTCCATCTTCTGTGGATTACCTGTGTTGACAAA 1332
QY 378 ValSerCysTyrValSerAspGlyAlaSerMetLeuThrPheGluSerLeuSerGlu 397
DB 1333 GTGTCATGCTATGTTCTGATGATGGTTTCAGCTATGCTGCTTTTGTAGTCTCTCTCAGAA 1392
QY 398 ThrAlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArg 417
DB 1393 ACCGAGAAATTTGCTAGAAAGTGGTTCCTTTTGTAAAGAGCAAAATTTGAACCAAGA 1452
QY 418 AlaProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThr 437
DB 1453 GCTCAGAAATTTTACTTTGCTCAAAAATAGATTACCTGAAGAGCAAAATTTCAACTTCA 1512
QY 438 PheValGlnGluArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsn 457
DB 1513 TTTGTTAAGAAAGAGCGCAATGAAGAGGAGTATGAAGAAATTCAAAGTGAAGATCAAT 1572
QY 458 AlaLeuValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThr 477
DB 1573 GCCCTTGTGCAAGACAGAAAGTGCCTGAAGAGGGGTGGACCATCGCTGATGAACT 1632
QY 478 ProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHis 497
DB 1633 GCATGGCTGGGAATATCTTAGGACCATCTGGCATGATTCAGGTTTTCTTGGGGCAC 1692
QY 498 SerGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlu 517
DB 1693 AGTGGTGGCTCGACACTGATGGAATAGATTACCAAGTCTTGTCTATGTCTCTCGTGA 1752
QY 518 LysArgProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgVal 537
DB 1753 AAGAGACAGGCTTTCAGCATCACAAAGAGCTGGTGAATGAATGCGCTGATTCGTGTA 1812
QY 538 SerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsn 557
DB 1813 TCTGCTGTGTCAGAAATGGTGCCTATCTCTCAATGGGATTGGACCATTTACTTCAT 1872
QY 558 AsnSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLys 577
DB 1873 AGCAGAAAGCTCTTAGAGAGCAATGTCTTCATGATGATCCGCTCTAGAGAGGAA 1932
QY 578 ValCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAla 597
DB 1933 ACTTGTATGTACAAATTTCCACAGAGATTTGATGGCATTTGCTTGCAGATCATGCT 1992
QY 598 AsnArgAsnThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyPro 617
DB 1993 ANTCCGAACATAGTATTTCTTTGATATCAATGAAGAGTCTGGATGGCATTCAGGGTCCA 2052
QY 618 ValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnPro--- 636
DB 2053 GTTTACGTGGGACAGAGTGTCTTCAATAGACAGGCTTTGTATGATGATCATGATCTGTT 2112
QY 637 ---ProLysGlyProLysArgProLysMetValThrCysAspCysCysProCysPheGly 655

```
Db      2113  TTGACTGAGCTGATCTGGAGCCAAACATTTGTTATTAAAGAGCTGCTGT-----GGT 2163
Qy      656  ArgLysLysArgLysHisAlaLys----- 663
Db      2164  AGAAGGAAGAAAAGAAACAGAGTTATATGATGATAGTCAAAGCCGTATTATTGAAGAGAACA 2223
Qy      664  -----AspGlyLeuProGluGlyThrAlaAspMetGly 674
Db      2224  GAATCTTCAGCTCCCATCTTCAATATGAAGACATCGAAGAGGGTATTGAA-----GGT 2277
Qy      675  ValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGln 694
Db      2278  TACGAGGATGAAGGTCAGTCTTATGTCCAGAGGAAATTGGAGAAACGCTTTGGTCAG 2337
Qy      695  SerAlaAlaPheValThrSerThrLeuMetGluGlyValProProSerSerSer 714
Db      2338  TCTCCTATTTTCATTTGTCATCCACCTTTATGACACAAGGTGGCATACCACCTTCAACAAAC 2397
Qy      715  ProAlaAlaLeuLysGluAlaIleHisValIleSerCysGlyThrGluAspLysThr 734
Db      2398  CCAGCTTCTCTACTAAAGGAAGCTATCCATGTCATCATCTAGTTGTGGATATGAGGACAAACT 2457
Qy      735  AspTrpGlyLeuGluLeuGlyTrpIleThrGlySerIleThrGluAspIleLeuThrGly 754
Db      2458  GAATGGGGAAGAGATTGGCTGGATCTATGTTCACTAAGGAGGATATTTCTGACTGGG 2517
Qy      755  PheLysMetHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAlaPhe 774
Db      2518  TTTAAATGTCATGCAAGGGCTGGCAATCAATCACTACTGTCATGCCACCAACGACCTTGTGTT 2577
Qy      775  LysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeu 794
Db      2578  AAGGTTCTGACCAATCAATCTTTCCGATCGTCTTAATCAGGTGCTCCGTTGGGCTCTT 2637
Qy      795  GlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGly 814
Db      2638  GGGTCAGTGGAAATCTCTTAGTAGACATTTGCTCTATCTGGTATGGTTAC--AATGGA 2694
Qy      815  AsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrIleTyrProPheThrSer 834
Db      2695  CGATTGAAGCTTTTGGAGAGGCTGGCTTACATCAACACTATTGTATATCAATCACATCC 2754
Qy      835  LeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIle 854
Db      2755  ATTCCGCTTATTGCTATGTGTGCTCCCGCTATCTCCCTCTTACCACATAAAATTATC 2814
Qy      855  MetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePhe 874
Db      2815  ATTCCTGAGATTAGCAATATGCTGGGATGTTCTTCATCTCTTTTCGCCCTCCATTTT 2874
Qy      875  AlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsn 894
Db      2875  GCCACTGGTATATTGGAGCTTAGATGGAGTGGTGTGGCATTTGAAGATTGGTGGAGAAAT 2934
Qy      895  GluGlnPheTrpValIleGlyGlyValSerAlaHisLeuPheAlaValGlnGlyLeu 914
Db      2935  GAGCAGTTTGGGTTATTGGTGGCACCTCTGCCCATCTCTTCGCAGTGTTCAGGGTCTG 2994
Qy      915  LeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAsp 934
Db      2995  CTGAAAGTTGGCTGGGATTGATACCAACTTCACAGTTACCTCAAGGCATCT--GAT 3051
Qy      935  GluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProPro 954
Db      3052  GAGGATGGCGCTTTGCTGAGCTATATGTGTTTCAAAGTGGACCCAGTTTGTCTATTCCTCCG 3111
Qy      955  ThrThrLeuLeuIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsn 974
Db      3112  ACCACTGTCTTGTATTAACTGGTGGATGGTGGCAGGAAATTTCTATGCCATTAAAC 3171
Qy      975  AsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIle 994
```

```
Db      3172  AGTGGCTACCAATCTCTGGGTCGCTCTTTTGGAAAGCTGTTCTTCGATCTGGGTGATC 3231
Qy      995  ValHisLeuTyrProPheLeuLysGlyLeuMetGlyValArgGlnAsnArgThrProThrVal 1014
Db      3232  CTCATCTCTTACCCCTTCTCTCAAGGGTCTCATGGGAAGGCAGAACCCGACACCAACATC 3291
Qy      1015  ValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAsp 1034
Db      3292  GTCATTTGCTGGTCCATCTCTTTCATCTATCTTCTCTTCTGCTGGGTGAAGATCGAT 3351
Qy      1035  ProPheIleValArgThr---LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db      3352  CCTTTTCATCTCCCGCACAGAAAGCTGCTGCTGGGCAATGTGGCGTCAACTGC 3408
```

Search completed: March 12, 2005, 14:31:17

Job time : 450 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	5648	100.0	3443	8	AY372246	AY372246 Cera maye
2	5347.5	94.7	3631	8	AK121170	AK121170 Oryza sat
3	4947	87.6	165909	8	AP005420	AP005420 Oryza sat
4	4947	87.6	187410	8	AP005579	AP005579 Oryza sat

[illegible]

```
/mol_type="mRNA"  
/db_xref="taxon:4577"  
1..3443  
/gene="CesA12"  
92..3250  
/gene="CesA12"  
/gene="CesA12"  
/note="CESA12; ZmCESA12"  
/codon_start=1  
/product="cellulose synthase catalytic subunit 12"  
/protein_id="AAR23312.1"  
/db_xref="GI:38532104"  
/translation="MEASGLVAGSHNRNBLVLRHEDPKPLRALSGQVCEICGDEV  
GLTVDFGLFVACNECGFPQPCVEYERREGTQPCQCKTRYKRLKSGPRVAGDDPEE  
DIDLEHEFINDEQOLEGNNMONSOITEAMLHGRMSYGRPGDDGDNNTGDDPEI  
ITGSRSPVSGEPEPTQKRGHGVSSSLKRIHPYVPVSEPSAKWDEKESVSKERMD  
DWKSKQILGGADPEMDADVALNDEARQPLSRKVSIASSKNPFRMVIWVRVLVLA  
FFURYLRLHPVDAIGLVLSIICEIWFALISWLDQPPKFPFIDIRETYLDRLSRYER  
EGFSLSLVDVSTVDPLKPLVNTLSILAVDVPDKVSCVSDGASMLTF  
ESLSETAFAKWKVPFCCKFGIIPRAPEFYSUKVDLKDQVQPTFFQERRAKREYE  
EPKVINALVAKAMKVAEGWIMKDGTPWPNTRDHPGMIQVFLHSGGHDTEGNELE  
PLRVYVSREKRPQGHKKAGMNAIJRVSAVLTNAPFNLNLDCHVINNSKAIAREAM  
CFLMDPOGRKVCIVQFPQPDGLDVHDIRANRNTVFEDINMKGLDGIQGPVYVGTGC  
VFRQALYGNPPKPGKRPKMWTCDCPCFRKKRKHAKDGLPEGTADMGMVSDKEML  
MSHNFERRQCSAFTSTLMEEGVPPSSPAALKKEAIIHVISCYEDKTDWGLLE  
GWYISITEDILTQKMHCRWRSVCMKRAAFKGSAPINLSDRLNOLVRWALGSVE  
IFSRHSPLLYGKGNLKWLERPAYINTTYPSTLPLLAYCTLPVCLLTGKFIIMP  
SISTFASLFTALPMSIATGILEMRWSGISIEMWNEQPFVWIGVSAHLFAVVOGL  
LKVLAGIDTNTVTSKATGDEDEFAELYAFKWTLLIPTTLLIINIVIGVAGISDA  
INNGYQSGWGLFGKLFPAFWVI VHLYPFLKGLMGRQNRTPTVVVISILLASIFSLW  
VRIDPFIVRTKGPDRVQCINC"
```

ORIGIN

```
Alignment Scores:  
Pred. No.: 0 Length: 3443  
Score: 5648.00 Matches: 1052  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0  
US-10-627-132-30 (1-1052) x AY372246 (1-3443)  
Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20  
Db 92 ATGGAGGCCACGCGCGGCTGGTGGCGGCTCGCACACCGAAGACGAGTGTCTGATC 151  
Qy 21 ArgGlyHisGluAspProLeuArgAlaLeuSerGlyGlnValCysGluIleCys 40  
Db 152 CGGGGCCACGAGGACCCCAAGCCGCTGCGGGCGCTGAGCGGGCAGGTGTGCGAGATGCG 211  
Qy 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60  
Db 212 GCGCAGAGGTGCGGCTCACGGTGGACGCGGACCTCTTCGTCGCTGCAACGAGTGGCGG 271  
Qy 61 PheProValCysArgProCysTyTGluTyArgArgGluGlyThrGlnAsnCysPro 80  
Db 272 TTCCTCCGTGTCGGCCCTGTACGAGTACGAGCGCGGAGGCGGACGCGAATGCGCCC 331  
Qy 81 GlnCysLysThrArgTyLysArgLeuLysGlySerProArgValAlaGlyAspAsp 100  
Db 332 CAGTGCAGACGCGCTACAAGCGCTCAAGGGAGCGCCGAGGTTCGCGGGGACGATGAC 391  
Qy 101 GluGluAspIleAspAspLeuHisGluPheAsnIleAspAspGluAsnGlnArg 120  
Db 392 GAGGAGGACATCGAGCCTCGAGCACGAGTTCACATCGACGACGAGAAATCAGCAGAGG 451  
Qy 121 GlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMet 140  
Db 452 CAGCTGGAGGGCAACATCGACAGACAGCAGATCATCCGAGGCGATGTCACGCGCAGGATG 511  
Qy 141 SerTyrglyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIleProProIle 160  
Db 512 AGCTACGGGAGGGGGCCCGACGACGCGGACGCGCAACACCCCGCAGATCCCGCCCATC 571
```

```
Qy 161 IleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrgly 180  
Db 572 ATCACCGGCTCCCGCTCCGTCGCGGTGAGCGGTGAGTTCCTCGATTACCAACGGGTATGCC 631  
Qy 181 HisGlyGluValSerSerSerLeuHisLysArgIleHisProTyroProValSerGluPro 200  
Db 632 CACGGCGAGGTCTCGTCTCTCCCTGACAAAGCGCATCCATCCGTACCCCTGTGTCTGAGCCA 691  
Qy 201 GlySerAlaLysTrpAspGluLysLysGluValSerTrpLysGluArgMetAspAspTrp 220  
Db 692 GGAGGTGCCAAGTGGGACGAGAAAGAAAGTGGCTGGAAGGAGAGGATGGACGATGG 751  
Qy 221 LysSerLysGlnGlyIleLeuGlyGlyValAspProGluAspMetAspAlaAspVal 240  
Db 752 AAGTCCAAAGCAGGCGCATCTCGCGGGCGCGCCGATCCCGAAGACATGGACGCGGACGTG 811  
Qy 241 AlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSerSerLys 260  
Db 812 GCACCTGAACGACGAGCGAGGCGCGCTCTCGAGGAAGGTGTCTGATCCGCTCGAGCAAG 871  
Qy 261 ValAsnProTyroArgMetValIleValValArgLeuValValLeuAlaPhePheLeuArg 280  
Db 872 GTGAACCCGTACCGGATGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 931  
Qy 281 TyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerIleIleCys 300  
Db 932 TACGATATCTGACCCCGTCCCGGACGATCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 991  
Qy 301 GluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsp 320  
Db 992 GAGATCTGGTTCGCCATCTCTGATCTCTCGACAGTTCGCCAAGTTCGCCAAGTTCGCCAAGTTC 1051  
Qy 321 ArgGluThrTyroLeuAspArgLeuSerLeuArgTyroGluArgGluGlyGluProSerLeu 340  
Db 1052 CGCGAGACGTACCTCGACCGGCTCTCCCTCAGGTACGAGAGGAGGAGGAGGAGGAGGAGGAGG 1111  
Qy 341 LeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluProProLeuVal 360  
Db 1112 CTGTGCGGCGTGGACCTGCTGCTGAGCACCGTGGACCCGCTCAAGGAGCGCGCGTGTGTG 1171  
Qy 361 ThrAlaAsnThrValLeuSerIleLeuAlaValAspTyroProValAspLysValSerCys 380  
Db 1172 ACCGCCAACACCGTGTCTCTCATCTCGCGGTAGACTACCCCGTGGCAAGGTCTCTCTGTC 1231  
Qy 381 TyrValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGlu 400  
Db 1232 TAGCTCTCCGACGACGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1291  
Qy 401 PheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArgAlaProGlu 420  
Db 1292 TTCGCGCGCAAGTGGGTGCCCTTCTGCAAGAAGTTCGCGCATCGAGCCCCCGCGCGGAG 1351  
Qy 421 PheTyroPheSerLeuLysValAspTyroLeuLysAspLysValGlnProThrPheValGln 440  
Db 1352 TTCTACTTCTCGCTCAAGGTGCGACTACCTCAAGGACAAAGGTGCGAGCCACCTTCGTGCGAG 1411  
Qy 441 GluArgArgAlaMetLysArgGluTyroGluGluPheLysValArgIleAsnAlaLeuVal 460  
Db 1412 GAGCGCGCGCCCATGAAGAGAGAGTATGAGGAGTTCAGAGGTTCGCGATCAAGCGCGCTGTGT 1471  
Qy 461 AlalysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrProTrpPro 480  
Db 1472 GCCAAGGCCATGAAGTTCGCGCAGAGGGGTGGATCATGAAGGACGCGCGCGCGCGCGCC 1531  
Qy 481 GlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGly 500  
Db 1532 GGGAAACAACACCCCGCAGCACCCCGCATGATCCAGGTGTCTCTGGGCGCACAGCGCGCGC 1591  
Qy 501 HisAspThrGluGlyAsnGluLeuProArgLeuValTyroValSerArgGluLysArgPro 520  
Db 1592 CACGACACCGAGGGCAACGAGTCCGCCCGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1651
```

Qy	521	GlyPheGlnHisHisValysAlaGlyAlaMetAenAlaLeuIleArgValSerAlaVal	540	Db	2732	ATCGCGTGGAGCGGGTGGAGCATCGAGGAGTGGTGAGAAACGAGCTTCTGGGTCTATC	2791
Db	1652	GAATTCACAGCACCAAGAGCCCGGCCATGAACGCTCTGATTTCGGCTCTCCGCCGTG	1711	Qy	901	GlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAlaGly	920
Qy	541	LeuThrAenAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLys	560	Db	2792	GGCGCGTGTCCGGCGCATCTCTTCGCGCTCGTGCAGGCGCTCTCAAGGTCTCTCCCGGG	2851
Db	1712	CTGACCAACGCGCCATTATGCTCAACTTGGACTGTGATCACTACATCAACACAGCAAG	1771	Qy	921	IleAspThrAenPheThrValThrSerLysAlaThrGlyAspGluAspGluPheAla	940
Qy	561	AlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysValCysTyr	580	Db	2852	ATCGACACCAACTTTCACGTCACCTCAAGGCCACCGCGCAGAGGACGAGTTCGCC	2911
Db	1772	GCCATTCGGGAGGCCATGTCTTCATGGACCTCAGGTCCGCGGGAAGTCTGTCTAC	1831	Qy	941	GluLeuTyrAlaPheLysTyrThrThrLeuLeuIleProProThrThrLeuLeuIleIle	960
Qy	581	ValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsnArgAsn	600	Db	2912	GAGCTCTACGCTTCAAGTGACCAACGCTCTCTCATCCCGCCACCGCTGTCTCATCAT	2971
Db	1832	GTTTCAGTTCGCGAGAGTTTCACGCGCATCGACGTGCAGCGCATCCGTAACAGGAAC	1891	Qy	961	AsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTrp	980
Qy	601	ThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProValTyrVal	620	Db	2972	AACGTCATCGCGGTCTGGCGCGCATCTCCGACGCCATCAACACGGGTACGAGTCTGG	3031
Db	1892	ACCGTCTTCTTCGACATCAACATGAAGGGCTGGACGGCATCCAAAGGCCCGGTGTACGTC	1951	Qy	981	GlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPhe	1000
Qy	621	GlyThrGlyCysValPheArgGlnAlaLeuTyrGlyTyrAsnProProLysGlyPro	640	Db	3032	GGGCGCCCTTTCGGCAAGCTCTTCTTCGCTTCTGGGTCTATCGTCCACCTCTACCGTTC	3091
Db	1952	GGGACAGAGTGTTCGGCGCCAGCGGCTCTACGGCTACAAACCTCCCAAGGGACCC	2011	Qy	1001	LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIleTrpSerIle	1020
Qy	641	LysArgProLysMetValThrCysAspCysCysProCysPheGlyArgLysValArgLys	660	Db	3092	CTCAAGGGGCTCATGGGCGCCAGAACAGGACGCCACCGTGTGTGTCTATCTGGTCCATT	3151
Db	2012	AAGAGGCCAAGATGTGACCTGCGACTGCTGCGCGTCTTCGGCGCGCAAGACGGAAA	2071	Qy	1021	LeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThr	1040
Qy	661	HisAlaLysAspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAspLysGlu	680	Db	3152	CTGCTGGCTCCATCTTCT	3211
Db	2072	CACGCCAAGGACGGCTCCGGAGGGCACCGCTGATATGGGAGTAGATAGCGACNAGAG	2131	Qy	1041	LysGlyProAspValArgGlnCysGlyIleAsnCys	1052
Qy	681	MetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThr	700	Db	3212	AAGGGCGCGGACGTTCAGGACGTGGGCATCAATTGC	3247
Db	2132	ATGCTCATGTCCCATGAACTTCGAGAAGCGTTCGGCGAGTCCGCGCGTTCGTCAAG	2191	RESULT 2			
Qy	701	SerThrLeuMetGluGluGlyGlyValProProSerSerSerProAlaAlaLeuLys	720	AK121170	3631 bp	mRNA	linear
Db	2192	TCGACGCTGATGGAGAGCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2251	AK121170			PLN 29-OCT-2003
Qy	721	GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeu	740	LOCUS			
Db	2252	GAGGCCATCCATGTCATCAGCTGGCTACGAGGACAGACGACTGGGGCTGGAGCTG	2311	DEFINITION			
Qy	741	GlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg	760	ACCESSION			
Db	2312	GGGTGGATCTACGGGTCTGATCAGGAGGACATCTCTGACGGGTTCAAGATGACCTGCCGC	2371	AK121170.1	GI:37990793		
Qy	761	GlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIle	780	VERSION			
Db	2372	GGGTGGCGCTCCGTGTACTGTCATGCCGAGCGGGCGGGTTCAGAGGGTCCGCGCGATC	2431	KEYWORDS			
Qy	781	AsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePhe	800	SOURCE			
Db	2432	AATCTATCGGACCGTCTCAACAGGTGTCTCGGTGGGCGCTGGGGTCTCGATCTTC	2491	ORGANISM			
Qy	801	PheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGlu	820				
Db	2492	TTACGCGGCACAGCCCTCTGTGTACGGCTACAGGACGCGCAACCTCAAGTGGCTGGAG	2551				
Qy	821	ArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyr	840				
Db	2552	CGCTTCGCTCATCAACACACCATCTACCCCTTCACTCCGCTCCGCTCCGCTCCGCTAC	2611				
Qy	841	CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr	860				
Db	2612	TGCACCTTCCCGCGCTGCTCTCTACCGCGAAGTTCATCTGCTCCGCTCCGCTCCGCT	2671				
Qy	861	PheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGlu	880				
Db	2672	TTCCGACGCTTCTTCTCATCGCTCTTCTATGTCATCTTCGCGAGCGGATCTCTGGAG	2731				
Qy	881	MetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGlnPheTrpValIle	900				

AK121170 3631 bp mRNA linear PLN 29-OCT-2003
 Oryza sativa (japonica cultivar-group) cDNA clone: J023081B08, full insert sequence.

AK121170
 AK121170.1 GI:37990793
 FLI_CDNA; CAP trapper.
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
 1 The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Koima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Onitsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Otsu, Y., Tsunoda, Y., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN: Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Oka, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.
 Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

TITLE
 japonica rice
 JOURNAL
 Science 301 (5631), 376-379 (2003)
 MEDLINE
 22752273
 PUBMED
 12869764
 REFERENCE
 2
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,

Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiroaka, T., Hori, F., Hotta, I., Iida, Y., Ikeda, R., Inamura, K., Ikehata, K., Imokani, K., Iehibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Mateubara, K., Mateuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakahama, Y., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otsomo, Y., Ryu, R., Saichou, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shiehiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.

Collection, mapping, and annotation of 28K full-length cDNA clones from japonica rice

Unpublished

3 (bases 1 to 3631)

Kikuchi, S.

Alignment Scores:		
Pred. No.:	0	Length: 3631
Score:	5347.50	Matches: 992
Percent Similarity:	96.78%	Conservative: 30

873	TGAGCAGAGTGTAAACCGGTACCGGATGGTGATCATCTCGTCTCGTGGTGCTCGGGTTC	932
Db		
278	PheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTyrLeuValSer	297
Qy		
933	TTCTCTGGGTACCGCATCTCCACCCGGTGCCGACGCCATCCCGTGTGGCTCACACCTCC	992
Db		
298	IleIleCysGluIleTyrPheAlaIleSerTyrPheLeuAspGlnPheProLysTyrPhe	317
Qy		
993	ATCATCTCGAGATCTGTTGCGCGTGTCTGGATCTCTGACCAAGTCTCCCAAGTGGTAC	1052
Db		
318	ProIleAspArgGluThrTyrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlu	337
Qy		
1053	CCGATCGACAGGAGAGACTTACCTCGACCGCTCTCCCTCGCTACGACGCGAGGGGAG	1112
Db		

Query Match:	93.94%	Mismatches:	29
DB:	8	Indels:	5
		Gaps:	4
US-10-627-132-30 (1-1052) x AK121170 (1-3631)			
QY	1	MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle	20
DB	96	ATGAGGCGAGGCGCGGCTGTCGCGGTCGCACACCGGAACGAGTGTGTGTCGATC	155
QY	21	ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGluIleCys	40
DB	156	CGGGGGCACGAGGAGCCCAAGCCGCTGCGGGCGCTGAGCGGCGAGGTGTGCGAGATATGC	215
QY	41	GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly	60
DB	216	GCGGACGAGTGGCGGCACCGTCGACGGCGACCTCTTCGTCGCTGCCAACGAGTCGCGC	275
QY	61	PheProValCysArgProCysTyrGluTyrGluAArgGluGlyThrGlnAsnCysPro	80
DB	276	TTCCCGGTGCGCGCCCTGCTACGAGTACGAGCGCGCGAGGCGACCCAGAACTGCCCC	335
QY	81	GlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAspAsp	100
DB	336	CAGTGCACACCGCTTACAAGCGCCTCAAGGGGAGCCCGAGGTGCCAGGGGACGAGGAC	395
QY	101	GluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnIn---	119
DB	396	GAGGAGGACATTGACGACCTGGAGCAGGATTCACATCGACGACGAGGAGGAGGAG	455
QY	120	---ArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly	138
DB	456	CTGCAGCAGGATCAGGATGGCATGCAGAACAGCACCATCACCGAGGCGATGCTGCACGCG	515
QY	139	ArgMetSerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIlePro	158
DB	516	AAGATGAGCTACGGGAGGGGCCCCGACGACGCGCGGCAACAGCACCCCG---CTCCCG	572
QY	159	ProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGly	178
DB	573	CGATCATCACCGCGCTCGCTCCGTCGCGTGAGCGGGAGTTCCCGATATCGAACAGC	632
QY	179	TyrGlyHisGlyGluValSerSerSerLeuHisLysArgIleHisProTyrProValSer	198
DB	633	CATGGCCATGGCGAGTTCTCTCTCTCCCTGCACAGCGCATCCACCCCTACCCGGTGTCT	692
QY	199	GluProGlySerAlaLysTrpAspGluLysLysGluValSerTrpLysGluArgMetAsp	218
DB	693	GAGCCAGGAGTGCAAGTGGGACGACGAGAAGAAGAGGTGAGCTGGAAGGAGAGATGGAC	752
QY	219	AspTrpLysSerLysGlnGlyIleLeuGlyGlyAla---AspProGluAspMetAsp	237
DB	753	GACTGGAATCCCAAGCAGGECATCGTCGCGCGCGGCCCCCGATCCGACGACTAGCAC	812
QY	238	AlaAspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAla	257
DB	813	GCCGACGTCCCACTGAACGACGAGGCGAGSCAGCCGCTGTCGAGGAAGGTGTCGATCGCG	872
QY	258	SerSerLysValAsnProTyrArgMetValIleValArgLeuValValLeuAlaPhe	277
DB	873	TCGAGCAAGGTGAACCGTACCGGATCGGTGATCATCTCCGTCTCGTGTGTCGCGGCTTC	932
QY	278	PheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSer	297
DB	933	TTCCCTCCGTTACCGCATCTCTCACCGCGGTGCCGACGCCATCCCGCTGTGGCTCACCTCC	992
QY	298	IleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPhe	317
DB	993	ATCATCTCGAGATCTGTTGTCGCGGTGTCGTGATCTCTCGACCCAGTTCGCCCAAGTGTAC	1052
QY	318	ProIleAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGlu	337
DB	1053	CCGATCGACAGGAGACCTACTCTGACCGGCTCTCCCTCCGCTACGACGCGAGGGGAG	1112

Qy	338	ProSerLeuLeuSerAlaValAspLeuPheValSerThrValAspProLeuLeysGluPro	357
Db	1113	CCGTCGCTGCTGCTCGGCGGTGGACCTGTTCGTGTACGACGCGTGGATTCGGCTCAAGAGCCG	1172
Qy	358	ProLeuValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLys	377
Db	1173	CCGCTGGTCAGCGCCACACCGTCTGTCCATCTTCGCCGTCGACTACCCCGTCGCAAG	1232
Qy	378	ValSerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGlu	397
Db	1233	GTGTCTCTGCTACGTCCTCCGACGCGCGCTCCATGCTCACGTTCCGAGTCCGTGTCGGAG	1292
Qy	398	ThrAlaGluPheAlaAArgLysTrrValProPheCysLysLysPheGlyIleGluProArg	417
Db	1293	ACGCGCGAGTTCGCGCCGCAAGTGGTGCCATTCTGCAAGAAGTTCACGATCGAGCCCCGC	1352
Qy	418	AlaProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThr	437
Db	1353	GCCCGGAGTTCTACTTCTCCAGAAAGTGCATCTACCTCAAGCAAGGTCCATCCCAAC	1412
Qy	438	PheValGlnGluArgAlaMetLysAArgGluTyrGluGluPheLysValAArgIleAsn	457
Db	1413	TTTCGTCCAGGAGCGCGCCGTCATGAAGAGAGTACGAGGAGTTCAAGGTGAGGATCAAC	1472
Qy	458	AlaLeuValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThr	477
Db	1473	CGCTGTGTGGCAAGGCGCAAGAGGTCCCGCGAAGGGTGGATCATGAAGAAGCGGACG	1532
Qy	478	ProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHis	497
Db	1533	CCATGGCGCGGGAACAACAACCCGGCGACACCCGGGCGATGATCCAGGTGTTCTTCGCGCAC	1592
Qy	498	SerGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlu	517
Db	1593	AGCGCGCGCCACACACCGAGGGCAACGAGCTCCCGCGCTCGTCTACGTCCTCCCGTGAG	1652
Qy	518	LysArgProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAlaLeuIleAArgVal	537
Db	1653	AAGCGCCCCGGCTTCCAGCACCAACAAGAAGCGCGGCCCATGAACGCCCTCATTTCTGTTGTG	1712
Qy	538	SerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsn	557
Db	1713	TCGCGCGTGTGACGAACGCGCCGTTTCATGCTCDACTTGGATTTCGATCACTACATCAAC	1772
Qy	558	AsnSerLysAlaIleAArgGluAlaMetCysPheLeuMetAspProGlnValGlyAArgLys	577
Db	1773	AACAGCAAGGCCATCAGGGAGGGGATGTGCTTCTCATGGATCCGAGGTCCGACGGAAG	1832
Qy	578	ValCysTyrValGlnPheProGlnAArgPheAspGlyIleAspValHisAspAArgTyrAla	597
Db	1833	GTTTGCTACGTGCGATTCCCGCAGAGGTTTCAGCGGCATCCAGCTCCACGACCATACGCC	1892
Qy	598	AsnArgAsnThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyPro	617
Db	1893	AACCGCAACACCGTCTCTTCGACATCAAATGAAAGGGGCTTGATGGGATCCAGAGGCCG	1952
Qy	618	ValTyrValGlyThrGlyCysValPheAArgGlnAlaLeuTyrGlyTyrAsnProPro	637
Db	1953	GTGTACGTCGGGACAGGGTTCGTGTTTCAGGCGGCAGCGCTGTACGAGATCAACACCCACC	2012
Qy	638	LysGlyProLysArgProLysMetValThrCysAspCysCysProCysPheGlyAArgLys	657
Db	2013	AAGGGACCCAAAGAGGCCAAGATGGTGACTCTGGACTGTGCTGCCCTTGTCTTCGGGAGAAG	2072
Qy	658	LysAArgLysHisAlaLysAspGlyLeuProGluGly---ThrAlaAspMetGlyValAsp	676
Db	2073	AAGCGGAAGCACGCGAAGGACGCGCTCCCGAGGCGCGTCCCGCGCGCGCGGATGGAC	2132
Qy	677	SerAspLysGluMetLeuMetSerHisMetAsnPheGluLysAArgPheGlyGlnSerAla	696
Db	2133	AGGCACAAGGAGATGCTCATGTGCGAGATGAATCTCGAGAAGCGGTTTCGGGCACTCGCG	2192

Qy	697	AlaPheValThrSerThrLeuMetGluGluGlyValProProSerSerProAla	716
Db	2193	GCCTTCGTACGTCGACGGCTGATGGAGGAAGCGCGCTCCCGCGTGTCTCAAGCCCGCGC	2252
Qy	717	AlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrp	736
Db	2253	GCCTCTCTCAAGGAGGCCATCATCATGTCATCAGCTGGGCTACGAGGACAAGACCGACTGG	2312
Qy	737	GlyLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLys	756
Db	2313	GGTCTCGAGCTGGGGTGGATCTACCGGTGCATCACGGAGGACATCTCTAACGGGGTTCAG	2372
Qy	757	MetHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGly	776
Db	2373	ATGCACCTGCCCGGGTGGAGGTCGGTGTACTGCATGCCGAAGAGGCGCGCTTCAAGGGG	2432
Qy	777	SerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySer	796
Db	2433	TCAGCGCGCATCAACCTATCTGACCGTCTCAACACGAGGTGCTCCGGTGGCGCTCGGCTCC	2492
Qy	797	ValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyValAsnLeu	816
Db	2493	GTTCGAGATCTTCTTCAGCGCGGCACAGCCCGCTCTCTACGGCTACAGAACCGAACCTC	2552
Qy	817	LysTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuPro	836
Db	2553	AACTGGCTCGAGCGCTTCTCTACATCAACACACCATCTACCCCTTCTCACTTCTCTCCCC	2612
Qy	837	LeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetPro	856
Db	2613	CTCCTCGCCTACTGCAACCTCACCCGCGCTGCTCCTCACCGGCAAGTTCATCATGCT	2672
Qy	857	SerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThr	876
Db	2673	CCGATTAGACAGTTTCGGAGTTGTCTTCATCGCGCTCTTCATCTCACTTCGCGACG	2732
Qy	877	GlyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGln	896
Db	2733	GGCATCTCTGGAGATGAGTGGAGCGGGTGGACATCGAGGAGTGGTGGAGGAACGAGCAG	2792
Qy	897	PheTrpValIleGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLys	916
Db	2793	TTCCTGGGTTCATCGCGCGCGTGTTCGGCGCACCTGTTCGGCGGTGGTCAGGCGCTGCTCAAG	2852
Qy	917	ValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAsp	936
Db	2853	GTCTGGCGGGATCGACACCACTTCACCGTCAGTCGAGGCCACCGAGACGAGGAC	2912
Qy	937	AspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThr	956
Db	2913	GACGAGTTCGGGAGCTCTACGCCCTTCAAGTGGACCACTCTCTCATCCCGCCCAACGAG	2972
Qy	957	LeuLeuIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnGly	976
Db	2973	CTGCTCATCTCAACATCATCGCGCTCGTGGCGCGTCTCCGACGCCCATCAACAACGCG	3032
Qy	977	TyrGlnSerTrpGlyProLeuPheGlyLysLeuPheAlaPheTrpValIleValHis	996
Db	3033	TCGAGGGCTGGGGCCGCTCTTCGGGAAGCTTCTTCGCGCTTCGGGTTCATCGTCCAC	3092
Qy	997	LeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValVal	1016
Db	3093	CTCTACCCCTTCTCAAGGGGGTCTATGGGAGGCGAAGACCGGACGCCACCATGTGTTGTC	3152
Qy	1017	IleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPhe	1036
Db	3153	ATCTGTTCGTGTGCTGCCCTCCATCTTTCTTGTCTCTGGGTTCAGGATTGATCCCTTC	3212
Qy	1037	IleValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys	1052
Db	3213	ACCATCAAGCCAGCGGGCCCTACGCTCAGGCAAGTCCGCATCACTGC	3260


```

/gene="P0418B08.5"
/notes="probably inactive due to including stop codon(s) in
CDS
pseudogene, transposase"
/pseudo
complement(29005..36438)
/gene="P0418B08.6"
complement(join(29005..29109,29221..29352,29652..29818,
30448..30637,31719..31850,32464..32571,32702..32797,
32891..32998,33106..33278,34215..34590,35076..35219,
35314..35988,36106..36171,36271..36438))
/gene="P0418B08.6"
/notes="start and end point are not identified"
complement(join(29005..29109,29221..29352,29652..29818,
30448..30637,31719..31850,32464..32571,32702..32797,
32891..32998,33106..33278,34215..34590,35076..35219,
35314..35988,36106..36171,36271..36438))
/gene="P0418B08.6"
/notes="similar to Oryza sativa chromosome 7,
OSUNBA0061L20.19"
/codon_start=1
/product="hypothetical protein"
/db_xref="GI:50725886"
/translation="MADRTSSDDSCSNKSDNKVADSDYANSISEEETSDDKVI
TKDLVTLKDFSGSKRKRARTDKPVRKQKRTKLEVDQDKSGSAESPKRSRLN
ITYFNSIEGLSNQRRIENSAFGLNFORCAIPLSFVKWIASHTDVSQDIIVNG
RSIPDNTNFILIGIPVNGGLEIKNDNDAGKHFFHFGSTKPLISFPGTKLLSDKG
NKLESDVLCFVWGLPGLPNSDTHPSKYLEPLSDIKSSKKNWKFVYEWLMT
YIAFORSEKSEKTSKTLGCGHLLAINVLDKFDGRNIPIGPPRISVWKGGMK
EYKIDTECKTGKRLIDDIISTVCKWFTREHSDTYTGKNAQAARVLDVCVMSKAS
LFTQKIDTTREKSEKSGKSGKGNVAHNGNAPPLFQNHPPNTDVCQLPKTP
QTSKTAHKTNNNTSTSGAPPTDKASSKDFVCSQSNNSVASRTRAKRLHSAS
PLSVISHPKYSDDKIRKARQKPLKATEPINTOPDREESNKDALFVEPICTIPA
KKEVQPTKNLESSTFVIDIEGPDADITGHTDKTFFILVYNSSEEHNSQD
TQDSNDIPNKSTNQPECLHFDKSKNSASLGHFLSQTSGTNIDIEHSSNQ
SSQSPDIGNSPRIAQMRPNQAABEROYSMIRIISLNASAGNBYRNFYFPH
RIVHPILMDNVKVTWSLSKSLSPRGVVDYVLNAYAKKIANDQNKENBYRNFYFPH
RTSLIFPCLNWFVFTVDIKGHFVFLSDIYDENNKYHKIKQGLLIPGIAMWEEF
SYVEKFSKFDIQYPPITRQNG"
40242..41638
/gene="P0418B08.7"
join(<40242..40308,40655..40900,40996..41112,41239..41339,
41570..>41638)
/notes="start and end point are not identified"
join(40242..40308,40655..40900,40996..41112,41239..41339,
41570..41638)
/gene="P0418B08.7"
/notes="predicted by GeneMark.hmm etc."
/codon_start=1
/product="hypothetical protein"
/db_xref="GI:50725887"
/translation="MPATEELESEATATWELDEAGEKAWGGALPHPPHPKPTAAVTRR

```

Alignment Scores:

Pred. No.:	0	Length:	165909
Score:	4947.00	Matches:	988
Percent Similarity:	74.18%	Conservative:	32
Best Local Similarity:	71.85%	Mismatches:	31
Query Match:	87.59%	Indels:	328
DB:	8	Gaps:	13

US-10-627-132-30 (1-1052) x AP005420 (1-165909)

Qy	1	MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuile	20
Db	5951	ATGAGGCGAGCGCGGCTGGTGGCCGGTTCGACCAACCGGAACGAGCTGGTCTGATC	6010
Qy	21	ArgGlyHisGluAspProLeuArgAlaLeuSerGlyGlnValCysGluileCys	40
Db	6011	CGGGGGCAGGAGGAGCCCAAGCCCTGCGGGCGCTGAGCGGGCAGGTTGCGAGATATGC	6070

Qy	41	GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysgly	60
Db	6071	GGCGACGAGGTGCGCGCACCGTTCGACGGCGACCTCTTCGTGCGTCAACAGGTGCGGC	6130
Qy	61	PheProValCysArgProCysTyrgluTyrgluArgGluGlyThrGlnAsnCysPro	80
Db	6131	TTCCCGGTGTGCGCCCTCTGCTACGAGTACAGCGCGCGGAGGACCCAGAACTGCCCC	6190
Qy	81	GlnCysLysThrArgTyrglyLysArgLeuLys	90
Db	6191	CAGTGCAGAGCCCGCTACAGCGCTCAAGGGTAAAAAACAACACTCACATCACGCTACG	6250
Qy	90	-----	90
Db	6251	CCCTTGATACCGCATCCGAGGTTTCGTTGATTGATCTCTAATGGCGATGGTGGT	6310
Qy	91	-----GlySerProArgValAlaGlyAspAspAspGluGluAspIleAsp	105
Db	6311	GGTGGTTTTGTACAGGGAGCCCGAGGGTCCCCGGGACGAGGACGAGGACATTGAC	6370
Qy	106	AspLeuGluHisGluPheAsnIleAspAspGluAsnGlnGln	123
Db	6371	GACCTGGAGCACGAGTTCAACATCGACGACGAGGACGAGCAGCTGCACGAGGATCAG	6430
Qy	124	GlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMetSerTyrgly	143
Db	6431	GATGGCATGCAGAACAGCACCATCACCGAGCGGATGTGTCACGCAAGATGAGCTACGG	6490
Qy	144	ArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIleProProIleIleThrGly	163
Db	6491	AGGGCCCCCAGCAGCGGCGAGCGCAACAGCACCCCG---CTCCCGCCGATCATCACCGC	6547
Qy	164	SetArgSerValPro	168
Db	6548	GCTCGCTCCGT-CCCGGTACACACACACACCTCACCCACTCACCAAAATTTCTCTCC	6606
Qy	169	-----ValSerGly	171
Db	6607	TCTTCTCACCTCAACATGTTTCACGAAATGTTCTTTGTAAAAAATAATTCAGGTGAGCGG	6666
Qy	172	GluPheProIleThrAsnGlyTyrglyHisGlyGluValSerSerSerLeuHisArg	191
Db	6667	GAGTTCGGATATCGAACACGATGGCCATGGCGAGTTCTCTCTCTCTCCCTGCACAAAGC	6726
Qy	192	IleHisProTyProValSerGlu	199
Db	6727	ATCACCCCTTACCGGGTGTCTGACCCAGGTAGTATACGAATTCATACACTAATTAATC	6786
Qy	200	-----ProGln	201
Db	6787	AATGTTCTTGATGAACACACATTCATCTCAATCATCTACCGATGAATTCGACACAGG	6846
Qy	201	YSerAlaLysTrpAspGluLysGluValSerTrpLysGluArgMetAspAspTrpLys	221
Db	6847	GAGTGCAGAGTGGGACGAGGAAGAGGAGGTGAGCTGGAAGAGAGGATGGACGACTGAA	6906
Qy	221	sSerLysGlnGlyIleLeuGlyGlyValAla	240
Db	6907	ATCCAGCAGGCGATCGTCGCGCGCGCGCCCGCCGATCCCGACGACTACGACGCGCGAG	6966
Qy	240	1-----	240
Db	6967	CCCACGTGACGAATCTCAGCTGAGCAGCTTACACTGATCATCATGATGACAACTAGTATA	7026
Qy	241	-----AlaLeuAsnAspGluAlaArgGlnProLeuSe	251
Db	7027	TTGATCATGCTGAACCTCTCTGTGGCTTGCGAGG-AACGACGAGGCGGAGGCGCGCTGC	7085
Qy	251	ArgGlyValSerIleAlaSerSerLysValAsnProTyArgMetValIleValVal	271
Db	7086	GAGGAGGTGTGATCGGTGAGCAAGGTGAACCCGTACCGGATGGTATCATCTCTCCG	7145


```
Db 121158 |||||GACCTGGACGAGTTCAACATCGACGACGAGACGAGCAGCTGCAGCAGGATCAG 121217
Qy 124 GlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMetSerTyrGly 143
Db 121218 GATGGCATGCAGAACAGCCACATCATCCGAGGCGATGTCGACGCAAGATGAGTACGGG 121277
Qy 144 ArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIleProProIleIleThrGly 163
Db 121278 AGGGGCCCCGACGACGGGACGGCAACAGCACCCCG--CTCCCGCCGATCATCACCGCG 121334
Qy 164 SerArgSerValPro----- 168
Db 121335 GTCGCTCGGT-CCCGGTACACACACACACCTCACCCCTCACACCAAAATTCCTCTCC 121393
Qy 169 -----ValSerGly 171
Db 121394 TCTTCTCACCTCAACATGTTACGAAATGTTCTTTGTTAAAAAAATTCAGGTGAGCGG 121453
Qy 172 GluPheProIleThrAsnGlyTyrGlyHisGlyGluValSerSerSerLeuHisIlybArg 191
Db 121454 GAGTTCGGGATATCAACAGCAGCATGGCCATGGCGAGTTCTCTCTTCCCTGCGACAGCGC 121513
Qy 192 IleHisProTyrProValSerGlu----- 199
Db 121514 ATCCACCCCTACCCGGTGTCTGAGCCAGGTAGTATAACGAATTCATCACATAAATATC 121573
Qy 200 -----ProGln 201
Db 121574 AATGTTCTTGATGAACACACACATTTGATCATCTCAATCATCTACCGATGAATTCGACCAG 121633
Qy 201 ySerAlaIlybTrpAspGluIlybGluValSerTrpIlybGluAArgMetAspAspTrpIly 221
Db 121634 GAGTCAAAAGTGGGACGAGAGAAGAGGTGAGCTGGAAGGAGGAGATGAGACGACTGGAA 121693
Qy 221 sSerIlybGlnGlyIleLeuGlyGlyAla---AspProGluAspMetAspAlaAspVa 240
Db 121694 ATCCAAGCAGGCGATCGTGGCGCGCGCCCGATCCCGACGACTACGACGCGCAGCT 121753
Qy 240 1----- 240
Db 121754 CCCACTGTAGCAATCTCAGCTGAGCAGCTTACACTGATCATGTCAGACAACTAGTATA 121813
Qy 241 -----AlaLeuAsnAspGluAlaArgGlnProLeuSe 251
Db 121814 TTGATCATGCTGAATCTCTCTGGGTTTCAGG-AAACGACGAGGCGAGCAGCCGCTGTC 121872
Qy 251 rArgIlybValSerIleAlaSerSerIlybValAsnProTyrArgMetValIleValValAr 271
Db 121873 GAGGAAGGTGTCGATCGCTCGAGCAAGGTGAACCCGTACCGGATGGTGATCATCTCTCG 121932
Qy 271 GluValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAspAlaIle 291
Db 121933 TCTCTGGTGTCTCGGCTCTCTCTCGGTACCGCATCTCCACCCCGGTGCCGACGCCAT 121992
Qy 291 eGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAs 311
Db 121993 CCCGCTGTGGCTCACCTCCATCATCTCGCAGATCTGGTTGCGCGGTGCTGGTATCTCGA 122052
Qy 311 pGlnPheProIlybTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSerLeuAr 331
Db 122053 CCAGTTCGCCAAGTGGTACCCGATCGACAGGAGACCTTACCTCGACCGGCTCTCCCTCG 122112
Qy 331 gTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSerThrVa 351
Db 122113 CTACGAGCGGAGGGGAGCCGCTGCTGCTGCGGGGTGGACCTGTTGTCGACGACCGGT 122172
Qy 351 lAspProLeuIlybGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAlaVa 371
Db 122173 GGATCCGCTCAAGGAGCGCGCTGGTCAGCGCCCAACACCGGTGCTGCCATCTCTCGCGGT 122232
Qy 371 lAspTyrProValAspIlybValSerCysTyrValSerAspAspGlyAlaSerMetLeuTh 391
```



```
QY 618 ValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProPro 637
Db 1827 GTGTATGTGGGACAGATCGTTCCTTCAAAAGGACAGCTTTGTATGCTATGACCTCC 1886
QY 638 LysGlyProLysArgProLysMetValThrCysAspCysCysProCysPheGlyArgLys 657
Db 1887 AAGGATCCAAAGCGCCCAAGATCGAAACCTGGACCTGCTGCCCATGTTTGGACGTCG 1946
QY 658 LysArgLysHisAlaLysAspGlyLeuProGluGlyThrAlaAspMetGlyValAspSer 677
Db 1947 AAAAAGAAAGATGCTAAGACTGGTGGCAGTTGTAGAGGAATGGAT-----ATAAT 1997
QY 678 AspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAa 697
Db 1998 GACAGAGGCTGTGTATGCTCCCATCAATGAAATTTGAAAGAAAGTTTGACCAATCAGCAAT 2057
QY 698 PheValThrSerThrLeuMetGluGluGlyGlyValProProSerSerProAlaAa 717
Db 2058 TTGCTAACTTCAACTTTAATGGAAGAGGTGGTGTACCTCTCTCGAGTCGGCAGCT 2117
QY 718 LeuLeuLysGluAlaLysHisValLysSerCysGlyTyrGluAspLysThrAspTropGly 737
Db 2118 CTGCTAAAGAGAGCCATCCATGTGATCAGTTGTGGATATGAAGACAAACCTGAATGGGA 2177
QY 738 LeuGluLeuGlyTropLysTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMet 757
Db 2178 CTGAGCTGGCTGGATTTACGGTTCATCAAGGAGATATCTCGACAGTTTAAAGATG 2237
QY 758 HisCysArgGlyTropArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySer 777
Db 2238 CATTTGCTGGCTGGAGTCTATTACTATGATGCAAGAGAGCTGCATTTAAGGGTTCA 2297
QY 778 AlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTropAlaLeuGlySerVal 797
Db 2298 GCTCCCATCAATCTATCAGATCGGCTAAACCAAGTCTCGCTGGGCTCTTGATCTGTT 2357
QY 798 GluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLys 817
Db 2358 GAAATTTTTTCAGTGTGTCACAGCCCTAATGTTGATGCTACAAAGAAAGGAAGCTCAAG 2417
QY 818 TrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeu 837
Db 2418 TGGCTCGAGAGTTTGGGTATGTGAACACAACTATCTACCCCTTCACCTCTCTTAGCACTC 2477
QY 838 LeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSer 857
Db 2478 GTTGCTACTGTTCCTCCCTCGCCATCTGCCCTGCTTACTGATAAATTTATCATGCCCCGAG 2537
QY 858 IleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGly 877
Db 2538 ATAAGCACTTTGGACGTCTTTCTTCACTGCTTTGTTTGTCAATCTTTTCCACGGGC 2597
QY 878 IleLeuGluMetArgTropSerGlyValSerIleGluTropTropArgAsnGluInPhe 897
Db 2598 ATTCTTGAGCTCAGATGAGCGGAGTAAGCATTTGAGGAATGGTGGAGAAACGAGCAATTC 2657
QY 898 TrpValIleGlyValSerAlaHisLeuPheAlaValGlnGlyLeuLeuLysVal 917
Db 2658 TGGGTTATAGTGTGTGTCTGCTCAGCTCTTTGCTGTGTGTTGTTGTTGTTGTTGTTGTTGTT 2717
QY 918 LeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspAsp 937
Db 2718 TTAGCAGGTATCGACCTAAACTTCACTGTCATCTCAAGGCTACA-----GACGATGAC 2771
QY 938 GluPheAlaGluLeuTyrAlaPheLysTropThrThrLeuLeuIleProProThrThrLeu 957
Db 2772 GATTTTGGAGAGCTTTATGCTTAAATGGCAACCCCTGCTTATCTCCCTCAACCACTATC 2831
QY 958 LeuIleIleAsnValIleGlyValAlaGlyLysSerAspAlaIleAsnAsnGlyTyr 977
Db 2832 TTAATCATCAACCTTTGTTGGATGTTGCTGGAGTCTCAGATGCCATAAACAATGGGTAC 2891
```

```
QY 978 GlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeu 997
Db 2892 CAGTCATGGGACCTCTATTTCGGAAGCTCTTCTTTCGCTTCTGGGTGATTTGCCATCTC 2951
QY 998 TyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValValIle 1017
Db 2952 TACCCATCTCTCAAAAGTCTAATGGGAGGCAAAACAGGACACCCGACTATTGTGGTTATA 3011
QY 1018 TrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTyrValArgIleAspPropheIle 1037
Db 3012 TGGTCAGTGTCTGCTGCTCCATCTCTCTCTGCTTGGTCCGGATTGATCCATTTGTG 3071
QY 1038 ValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3072 ATGAAACACAGGAGACCTGACACCAAGCAATGTGGACTCAACTGT 3116

RESULT 6
BT013780
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 3414
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="132670R"
/tissue_type="fruit"
/note="TZBY89"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3414
Score: 4784.50 Matches: 880
Percent Similarity: 90.72% Conservative: 78
Best Local Similarity: 83.33% Mismatches: 19
Query Match: 84.71% Indels: 19
DB: 8 Gaps: 7

US-10-627-132-30 (1-1052) x BT013780 (1-3414)

QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
Db 66 ATGGAAGCTAGTCAGGCTCTAGTTGCTGGTTTCATCAACCGGAATGAACCTTGTGCTATT 125
QY 21 ArcGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGluIleCys 40
Db 126 CATGGCCATGAAGAGCATTAAGCCATTGAAGGATTTGAGTGGACACAGTTTGTGACATATGT 185
QY 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
Db 186 GGGGATGAATAGGCTTTAACAGTGGATGGTGTATCTTTTGTGGCTTCAATGAGTGTGGT 245
QY 61 PheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsnCysPro 80
Db 246 TTCCTGTGTGTCGGCTTCTGCTACGAGTATGAGAGAAAGGAAGAACTCAACAGTGTCCA 305
QY 81 GlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAspAsp 100
Db 306 CAGTGCAGACCAAGATCAAAACGCTCTTAAGGGAGTCCAAAGGGTGGCGGGAGATGATGAT 365
```


Db 2487 AAGTGGCTGAACGATTATCATACATTAACAACAACACTATTTACCCTTCTACTCTTTTACCT 2546
Qy 837 LeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLeuLysPheIleMetPro 856
Db 2547 CTCCTTGGCTACTGACATCTCCCTGCTGTCGTGTTGCTCACTGGAAAATTCATATATGCCA 2606
Qy 857 SerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThr 876
Db 2607 GAGATAAGCACGTTTAGCAAGTCTGCTTTTTCATTCCTCTCTATCTTTTACAACG 2666
Qy 877 GlyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGln 896
Db 2667 GGAATCTTGGCTGAAGTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2726
Qy 897 PheTrpValIleGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLys 916
Db 2727 TTTTGGGTATTGAGGTGATATCTGCCACCTTTTCTGCTAGTACAGGTCTTCTCAAG 2786
Qy 917 ValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAsp 936
Db 2787 ATTTTGGCTGGCATAGATACCACTTTTACTGTGACATCTTAAGGCAACAGATGATGAGGAC 2846
Qy 937 AspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuLysLeuProThrThr 956
Db 2847 -----TTCGGAGAGTTATATGCTTCAAGTGGACTACTCTCTTATTCTCTCAACAACA 2900
Qy 957 LeuLeuIleAlaValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGly 976
Db 2901 ATTTCTCATATCAACTTGGTGGGTGTAGCCGGAATCTCTGACGCTATCAACACGGA 2960
Qy 977 TyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHis 996
Db 2961 TACAATTCATGGGTGCCCTTTTCGGAAGAGCTTTTCTTGGCTTTTGGTCTATTGTTTCAT 3020
Qy 997 LeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValVal 1016
Db 3021 TTATATCATATTCCTCAAGGGTCTTTATGGCGAGGAGAGCAACAGGACTCAACACCATTT 3080
Qy 1017 IleTrpSerIleLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPhe 1036
Db 3081 ATATGGTCTATATCTTTAGCTTCATCTTTCTTCTGCTGCTGGTTCGAAATGATCCATT 3140
Qy 1037 IleValArgThrLysGlyProAspValArgGlnCysGlyIleLeuAsnCys 1052
Db 3141 GTGCTCAAAACCAAGGACAGATGTCAAGCGATGTGGAGTGAAGTGC 3188
RESULT 7
BT004543 LOCUS
DEFINITION Arabidopsis thaliana AT5g17420/T10B6_80 gene, complete cds.
ACCESSION BT004543
VERSION BT004543.1 GI:28416516
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 3081)
Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,
Onodera,C.S., Palm,C.J., Dale,J.M., Hsuani,V.W., Wong,C., Wu,H.C., Yamada,K.,
Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and
Ecker,J.R.
Arabidopsis ORF clones
Unpublished
2 (bases 1 to 3081)
Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,
Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Hsuani,V.W.,
Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K.,
Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and
Ecker,J.R.

TITLE
JOURNAL
COMMENT
Alignment Scores:
Pred. No.: 0 Length: 3081
Score: 4652.00 Matches: 853
Percent Similarity: 89.10% Conservative: 87
Best Local Similarity: 80.85% Mismatches: 83
Query Match: 82.37% Indels: 32
DB: 8 Gaps: 10
US-10-627-132-30 (1-1052) x BT004543 (1-3081)
The Salk, Stanford, PSEC (SSP) Consortium members constructed and
sequenced the PUNI (ORF) clones using the RAPL cDNAs: Cheuk,R.,
Chen,H., Kim,C.J., Shinn,P., Bowser,L., Chan,M.M., Chang,C.M.,
Dale,J.M., Hsuani,V.W., Jones,T., Karlin-Neumann,G., Lam,B.,
Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J.,
Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
Yamada,K., Yu,G., Yuan,S., Davis,R.W., Theologis,A., and
Ecker,J.R.
Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.
FEATURES
Location/Qualifiers
source
1..3081
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="U22199"
/ecotypes="Columbia"
/notes="This clone is in pUNI 51"
CDS
1..3081
/notes="cellulose synthase catalytic subunit (IRX3)"
/codon_start=1
/product="AT5g17420/T10B6_80"
/protein_id="AAO42789.1"
/db_xref="GI:28416517"
/translation="MEASGLVAGSHNRNELVVIHNEEPKPLKNDLQDQFCICGQDI
GLTVEGDLFVACNECPYERREGTQNCQCKYKRLGSRVDEGDESE
DIDIEVEFNIEHQDKHSAEAMLYGKMSYGRGPDENGPRPPVACGHSEFPV
GGYNGEHLKRVHPYPSSEAGSGWRNDMKLOHGNLGPEDDDPENGLEIDE
ARPLSKRVFIASSKINPRMIVARLVILAVFLRLNPNVHDALGLMTLSVCEIW
FAVSWILDQPPKRPFIERTYLDRLSRYEREGERPNMLAPVDVFTVDPKSPPLVT
SNVLSILAMDYVPEKISCVSDDGASMLTFESLSEAEFAKWVFPCKFSIEPRAP
EMVTLKDYLDQKHVPTFKERRAMKREYEEFKVINAQAKASVPLSGWIMQDGT
PWPNNTKDHPGMIOVFLHSGGFDVEGHELPRLVYVSRKPGFQHHKAGAMNALV
RVAGVLTNAPFMLNLCDDHYVNSKAVREANCMFLMDPQIGKVCYVQFPQFQIDTN
DRYANRNTVFFDINMGLDGTQGVTVGCVFQALYGFEPKPKRPMKSCGCC
PCGRRKKNKFSKNDMGDAAGBAEGDKHELMSEMPKFTFGSSSIFVTLMSE
GCVGSSPAVLKKEAIIHVSCEYEDKTEWELGWIYGSITDILTFQWGHGWRMS
IYCMKRPAPKPSAPINLDRNLQVRLWALGSEVIFFSRSLPWLGYKGLKWLRF
AVANTTTPSTIPLAYCTIPALCILLTKFIMPSTFASLFIPLFMSIIVTGILE
LRWSGVSEIENWRNEQFWIGSIAHLFAVQGLLKLKILAGIDTNFTVTSKATDDDFG
ELVAFKWTLLIPPTVLLINIVGVAGSDAINNGYQSGWGLFGKLFPSFWIVHLY
PFLKGLMGRQNRTPPTTIWVNSVLLASIFSLLLVRIDFPVLTKRTPDTSKCGINC"

Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
Db 1 ATGGAAGCTAGCGCGGCTTTGTCGCGGTTCTCATAAACCGTAATGAACTAGTCCTATT 60
Qy 21 ArgGlyHisGluAspProLeuArgAlaLeuSerGlyGlnValCysGluIleCys 40
Db 61 CACAACCATGAAGCCAAAGCCTCTGAAGAATCTAGATGGACAATCTCTGTGAGATATGT 120
Qy 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
Db 121 GGAGATCAGATTGGTTTAAACAGTAGAAGGAGACCTCTTCGTAGCTTGGCAATGAGTGTGGT 180
Qy 61 PheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsnCysPro 80
Db 181 TTTCCGGCGGTGAGACCTTGTCTATGATGACGAGAGAGAGAGAGGAAACCAAACTGTCTCT 240
Qy 81 GlnCysIleThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAspAsp 100
Db 241 CAGTGTAAAGACTCGTTACAGCGTCTCAGAGGAGCCCAAGAGTGGAGGAGATGAAGAC 300
Qy 101 GluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnArg 120
Db 301 GAAGAAGATATGATGATATGAGTATGATGAATTAATATCGAATCATGAACAAGAT----- 354
Qy 121 GlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMet 140
Db 355 -----AAGCATAAAGCATTTCTGCTGAGGCTATGCTTTATGGGAAATG 396
Qy 141 SerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIleProProfile 160
Db 397 AGCTATGGAAGAGGTCTCGAGATGATGAGATGGG-----AGATTCCCACTGTT 447
Qy 161 IleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrGly 180
Db 448 ATAGCTGTGTGTCAT-----AGTGAGAGATTTTCAGTTGGAGGAGGTATGTT 495
Qy 181 HisGlyGluValSerSerSerLeuHisLysArgIleHisProTyrProValSerGluPro 200
Db 496 AATGGAGAA-----CATGGGCTTCATAAGGCGTGTGACCCCATATCCATCATCTTGAAGCT 549
Qy 201 GlySerAlaLysTyrAspGluLysGluValSerTyrLysGluValMetAspAspTyr 220
Db 550 GGGAGT-----GAGGAGAGATGCGCGGAAAGAAAGATGGATGACTGG 588
Qy 221 LysSerLysGlnGlyIleLeuGlyGlyAlaAspProGluAspMetAspAlaAspVal 240
Db 589 AAGCTCCAGCATGGAATCTT-----GGGCCAGAACCAAGATGAT--GATCCTGAGATG 639
Qy 241 AlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSerSerLys 260
Db 640 GGACTGATCGAGCGGACGCGCAGCCACTCTCGCGGAAAGTTCCCATTTGCCCTCAAGCAAG 699
Qy 261 ValAsnProTyrArgMetValIleValArgLeuValValLeuAlaPhePheLeuArg 280
Db 700 ATCAATCCATATCGATGGTTCATGCTGCTAGGCTTGGATTCAGACAGTTTTCGCGG 759
Qy 281 TyrArgIleLeuHisProValProAspAlaIleGlyLeuTyrProValSerIleIleCys 300
Db 760 TATAGGCTCTTGAATCCAGTCAGTATGCTCTGGGATATGCTGACCTCTGTGATCTGT 819
Qy 301 GluIleTyrPheAlaIleSerTrpIleLeuAspGlnPheProLysTyrTrpPheProIleAsp 320
Db 820 GAAATCTGGTTTCGCTGTCTCTTGGATTTGATCAGTTCCCGGAAAGTTCCCTATTGAA 879
Qy 321 ArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGluProSerLeu 340
Db 880 CGTGAGACTATCTAGATCGGCTTTCCTCAGGTACGAGAGAGAGGTAACCAAAATATG 939
Qy 341 LeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluProProLeuVal 360
Db 940 CTGCCCCCTGTAGATGCTTTGTGTCAGTACGCTGGACCATTTGAAGGAGCCTCCCTCGTC 999
Qy 361 ThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCys 380

Db 1000 ACATCCCAACACTGTGCTGCTCAATCTTGCCCATGACTACCCAGTTGAGAAAATCTCTCGC 1059
Qy 381 TyrValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGlu 400
Db 1060 TATGCTCTCTGACGAGCGGTCTCAATGCTTACATTCGAATCTCTCTCGAANAACCTGCTGAG 1119
Qy 401 PheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArgAlaProGlu 420
Db 1120 TTTGCAAGAAAATGGGTTCCCTCTGTAAGAAAATTTCTCATAGAGCCACCGGCAACCGGAG 1179
Qy 421 PheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThrPheValGln 440
Db 1180 ATGTACTTACGTTGAAAGTTGATTTATCTCCAGACAAAGTCCACCAACATTTGTTAAG 1239
Qy 441 GluArgArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAlaLeuVal 460
Db 1240 GAACGTCGAGCCATGAAGAGAGATATGAGGAGTTCAAGGTCAGGATCAACGCTCAAGTG 1299
Qy 461 AlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrProTyrPro 480
Db 1300 GCGAAGGCCTCAAAAGGTTCTCTAGAAAGTTGGATCATGCAAGATGGAACACCGTGGCA 1359
Qy 481 GlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGly 500
Db 1360 GGGAAACCAACCAAGGACCCCGGTATGATCCAAAGTCTTCTCGGCCACAGCGGAGGA 1419
Qy 501 HisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgPro 520
Db 1420 TTTGATGTCGNAAGGCGATGAGCTTCTCGGCTGTGTACGTGTCCTCGTGAAGCGTCT 1479
Qy 521 GlyPheGlnHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSerAlaVal 540
Db 1480 GGTTTTCAACACCAAGGAGTGGCGCATGAATGATCCCTGTTGAGTGGCAGCGCGTA 1539
Qy 541 LeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLys 560
Db 1540 CTCACAAATGCTCTCTTTCATGCTGAACTTGGACTGTGATCATCTATGATGATAAACCAAGCAAG 1599
Qy 561 AlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysValCysTyr 580
Db 1600 GCCGTGAGGAGCAAGATGTTTTCATGGATCTCTCAGATTTGGAAGAAGGTCTGCTAT 1659
Qy 581 ValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsnArgAsn 600
Db 1660 GTTCAGTTCTCCCTCAAGGTTTGTATGTCATTCACAAACGATCGTTACGCCACAGAAAC 1719
Qy 601 ThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProValTyrVal 620
Db 1720 ACAGTCTTCTTTGATATCAATATGAAAGGTCCTAGATGAAATCCAAAGGTCCTAGCTTTACGTT 1779
Qy 621 GlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProProLysGlyPro 640
Db 1780 GGTACTGTTGTTGTTTCAACACGACAGCTCTGTATGTTATGAACCAACCAAGGTCCT 1839
Qy 641 LysArgProLysMetValThrCysAspCysProCysPheGly-----ArgLysLys 658
Db 1840 AAACGTCCAAAGATGATAAGCTGTGTTGTTGCTTGTCTTGGCGCGCGGAGAAAGAAT 1899
Qy 659 ArgLysHisAlaLysAspGlyLeuProGluGlyThrAlaAspMet---GlyValAspSer 677
Db 1900 AAGAAATTTTCCAAAGAATGATGAATGGTGACCTAGCAGCCCTTGGAGGAGCAAGAGGT 1959
Qy 678 AspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAla 697
Db 1960 GATNAGAGGCAATTTGATGCTGGAATGACATTTGAGAAAACATTTGGCAATCATCCATC 2019
Qy 698 PheValThrSerThrLeuMetGluGluGlyValProProSerSerSerProAlaAla 717
Db 2020 TTTGTAACCTCAACTTTGATGGAAGAAGTGTGTTCTCTCGTCTCATCAAGTCTCGCAGTG 2079
Qy 718 LeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTyrGly 737

Db 2080 CTCCTTAAAGAGGCAATCCATGTCATAAAGCTCGCGTTATGAAGACAGACTGAGTGGGA 2139

Qy 738 LeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLeuMet 757

Db 2140 ACTGAGCTGGTGGATCTATGCTCTATACAGAGGATATTTTGCAGGATTCAGATG 2199

Qy 758 HisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySer 777

Db 2200 CATTCGCGTGGATGGAGGTCATATTACTGCTAAGAGGCGCTGCATTCAAAGGTTCA 2259

Qy 778 AlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerVal 797

Db 2260 GCTCCTATTATCTATCAGACAGGTTAAACAGCGTTTTCGCTGGGCACATTGATCGTT 2319

Qy 798 GluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLys 817

Db 2320 GAGATATTTTACCGCGGACAGTCTCTCTGTATGGCTACAAAGGAGCAAACTCAAG 2379

Qy 818 TrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeu 837

Db 2380 TGGCTTGAGCGTTTTCCTTATGCGCAACACAACTCTACCCCTTCACATCTATACCATT 2439

Qy 838 LeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSer 857

Db 2440 CTTGCTTACTGTATCTCTCCAGCCATCTGCTCTTACTGACAAATTCATGCCACCG 2499

Qy 858 IleSerThrPheAlaSerLeuPheIleAlaLeuPheMetSerIlePheAlaThrGly 877

Db 2500 ATAAGCACATTTGCTAGTCTCTTCTTCATCTCCTGTTATGTCATCTATTGACGGA 2559

Qy 878 IleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGluInPhe 897

Db 2560 ATCTTGGAATTGAGATGGAGCGAGTTAGCATTTGAGGAATGGTGGAACAGCAATTC 2619

Qy 898 TrpValIleGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysVal 917

Db 2620 TGGGTCTATGGAGAACTCTCAGCTCATCTCTTTCGCGTTGTCCAAAGTCTCTCAAAT 2679

Qy 918 LeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspAsp 937

Db 2680 TTAGCAGCAATGCACAACTTCACCGTCACATCAAGGCAACA-----GATGATGAT 2733

Qy 938 GluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeu 957

Db 2734 GACTTTGGAGAACTTTACGCATTCAAATGACAACACTGCTGATCCCTCAACAACATGTC 2793

Qy 958 LeuIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyr 977

Db 2794 TTAATCATAAACATTTGTGGCGTTGTTCAGGCACTCTCAGATGCCATTAACAATGGATAT 2853

Qy 978 GlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeu 997

Db 2854 CAGTCTTTGGGACCTCTATTTTGGTAAACTCTCTCTCTCTCTTTTGGGTCATGTTCATCTC 2913

Qy 998 TyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIle 1017

Db 2914 TACCAATTCCTCAAAGGTCGTGATGGGTAGACAGAACACACCAACCATTTGGTGATT 2973

Qy 1018 TrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIle 1037

Db 2974 TGGTCAGTGTATTGGCATCTATCTCTCTTTGCTTTGGGTAAAGAAATGTATCTCTTTGTG 3033

Qy 1038 ValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052

Db 3034 CTCGAACCAAGGACCTGACACTTCCAAAGTGTGGCATCAACTGC 3078

RESULT 8

AY139754

LOCUS AY139754 3355 bp mRNA linear PLN 02-SEP-2002

DEFINITION Arabidopsis thaliana AT5g17420/t10B6_80 mRNA, complete cds.

ACCESSION AY139754

VERSION AY139754.1 GI:22654964

KEYWORDS FLI_CDNA.

SOURCE
ORGANISM

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 3355)

REFERENCE
AUTHORS

Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Bowser, L., Carninci, P.,
Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y.,
Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,
Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,
Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C.,
Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W.,
Theologis, A. and Ecker, J.R.

Arabidopsis cDNA clones

Unpublished

2 (bases 1 to 3355)

REFERENCE
AUTHORS

Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Bowser, L., Carninci, P.,
Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y.,
Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,
Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,
Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C.,
Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W.,
Theologis, A. and Ecker, J.R.

Direct Submission

TITLE
JOURNAL

Submitted (09-AUG-2002) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PCEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Shinn, P., Chen, H.,
Cheuk, R., Kim, C.J., Bowser, L., Chan, M.M., Chang, C.M., Dale, J.M.,
Deng, J.M., Hsuan, V.W., Jones, T., Karlin-Neumann, G., Lam, B.,
Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Palm, C.J., Quach, H.L.,
Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C.,
Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and
Ecker, J.R.

Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES

source

1. .3355

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="5"

/clone="RAFL09-35-F05 (R22199)"

/ecotype="Columbia"

46. .3126

/note="cellulose synthase catalytic subunit (IRX3)"

/codon_start=1

/product="AT5g17420/t10B6_80"

/protein_id="AAM98075.1"

/db_xref="GI:22654964"

CDS

/translation="MEASAGLVASHNRNELVVIHNEBPKPLNLDGQFCEICGDOI
GLTVEGDLFVACNEGFPACPCYEYERREGTQNCPCQCKTRYKRLGSRVGEDEDE
DIDDIYEENIEHQDKHSAEAMLYGMSYGRGPDDENGFPVVIAGHSGSERPV
GGYNGEHLKRVHPYFSSAGSGGWRERMDWLQHLQNLGPPDDDDPENGLIDE
ARQPSRKVPFIASSKINPYRMVILAVFLRYRLNPDVHDLGLWLTSLVCEIWL
FAVSTLDGFPKFFPIETRYDLRLSRYERECEPNMLAPDVFVDTDLKLPPLVT
SNVLSILAMDYVKSVDGASMLTFBSLSETAFARWAKVFCFKFSIEPRAP
ENVFTLKVDLQDKVHTPVKERRAMKREYBPVKINAQWAKSVKPLGWTMQDQT
PFRGNTKXHPGMIQVPLHSGSPDVEGHELRLVYVSRKPGFQHHKAGAGNALV
RVAGLNTAVFFMLNLOCDHYVNSKAVRAMCFMLMDPQIGKKYVQVQFQRPFGIDTN
DRYANRNTVFFDINMKGLDGIQGVVYVGTGCVFKRQALYEPKPKPKRMISCGCC

ORIGIN PFLKGLMGRNQRPFTIIVIVWSVLLASIFSLWVRIDPFVLKTKGPDTSKCGINC"

Alignment Scores:
 Pred. No.: 0 Length: 3081
 Score: 4646.00 Matches: 852
 Percent Similarity: 89.00% Conservative: 87
 Best Local Similarity: 80.76% Mismatches: 84
 Query Match: 82.26% Indels: 32
 DB: 8 Gaps: 10

US-10-627-132-30 (1-1052) x AF088917 (1-3081)

Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
 Db 1 ATGGAAGCTAGCGCGGTCTTGTGCGCGGTCTCTAATACCGGTAATAGTGGTCAAT 60

Qy 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGluLeu 40
 Db 61 CACACCATTGAAGAGCCCAAGCCTCTGAAGAATCTAGATGGACAATCTCTGGAGATATGT 120

Qy 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
 Db 121 GGAGATCAGATTGTTTAAACAGTAGAAGAGACCTCTTCGTAGCTTCCATAGTGGTGGT 180

Qy 61 PheProValCysArgProCysTyrGluTyrGluArgGluGluThrGlnAsnCysPro 80
 Db 181 TTTCCGCGGTAGACCTTGTCTATGATGATGATGATGATGATGATGATGATGATGATGAT 240

Qy 81 GlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAspAsp 100
 Db 241 CAGTGTAAAGACTCGTTTAAACAGTAGAAGAGACCTCTTCGTAGCTTCCATAGTGGTGGT 300

Qy 101 GluGluAspLeuAspLeuGluHisGluPheAsnLeuAspGluAsnGlnArg 120
 Db 301 GAAGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 354

Qy 121 GlnLeuGluGlyAsnMetGlnAsnSerGlnLeuThrGluAlaMetLeuHisGlyArgMet 140
 Db 355 -----AAGCATAAAGCTATCTGCTGAGGCTATGCTTTATGCGAATAAG 396

Qy 141 SerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnLeuProPhe 160
 Db 397 AGCTATGGAAGAGGTCTCTGAGGATGATGAGAATGGG-----AGATCCACACCTGTT 447

Qy 161 IleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrGly 180
 Db 448 ATAGCTGTGTCTCAT-----AGTGGAGAAATTTCCAGTTGGAGGAGGTATGTT 495

Qy 181 HisGlyGluValSerSerSerLeuHisLysArgIleHisProTyrProValSerGluPro 200
 Db 496 AATGGAGAA-----CATGGGCTTCAATGAAGCGTGTGCACCATATCATCTGAAGCT 549

Qy 201 GlySerAlaLysTrpAspGluLysLysGluValSerTrpLysGluArgMetAspAspTrp 220
 Db 550 GGGAGT-----GAGGGAGGATGGCGGAAAGAAATGGAATGATGATGATGATGAT 588

Qy 221 LysSerLysGlnGlyLeuGlyGlyAlaAspProGluAspMetAspAlaAspVal 240
 Db 589 AAGCTCAGCATGGAATCTT-----GGCCAGAACCATGATGAT-----GATCTGAGATG 639

Qy 241 AlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSerSerLys 260
 Db 640 GGACTGATGACGAGGACGCGACCCACTCTCGCGGAAAGTTCCCATTTGCTCAAGCAAG 699

Qy 261 ValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPheLeuArg 280
 Db 700 ATCAATCCATATCGGATGGTTCATCGTTGCTAGGCTTGTGATCTTAGCAGTCTTTCTGCGG 759

Qy 281 TyrArgIleLeuHisProAspAlaIleGlyLeuTrpLeuValSerIleLeuCys 300
 Db 760 TATAGGCTCTTGAATCCAGTGCATGATGCTCTGGGATTTATGGCTGACCTCTGTGATCTGT 819

Qy 301 GluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsp 320
 Db 820 GAAATCTGGTTCGCTCTCTTGGATTTCTTGGATTCCTCCCAAGTGGTTCCTCATTTGAA 879

Qy 321 ArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGluProSerLeu 340
 Db 880 CGTGAGACCTATCTAGATCGGCTTTCCCTCAGGTACGAGAGAGAGTGAACCAATATG 939

Qy 341 LeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluProLeuVal 360
 Db 940 CTTGCCCTGTAGATGCTTTTGTAGTGTGAGCCCATTTGAAGAGAGCTCCCTCTGCTGTC 999

Qy 361 ThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCys 380
 Db 1000 ACATCCAAACACTGCTGCTCAATCTTGGCCATGAGCTACCCAGTTGAGAAAATCTCTCTGC 1059

Qy 381 TyrValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGlu 400
 Db 1060 TATGTTCTTGACACCGGTGCTTCAATGTTTACATTTCAATCTCTCTCGGAATCTGCTGAG 1119

Qy 401 PheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArgAlaProGlu 420
 Db 1120 TTTGCAAGAAATGGGTTCCTTCTGTAGAAATTTCTCCATAGACCCGCGCACCGGAG 1179

Qy 421 PheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThrPheValGln 440
 Db 1180 ATGTACTTTCAGTTGAAAGTTGATTTATCTCCAGGACAAAGTCCACCAACATTTGTTAAG 1239

Qy 441 GluArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAlaLeuVal 460
 Db 1240 GAACGTCGAGCCCATGAAGAGAAATATGAGGAGTTCAAGGTCAGGATCAACGCTCAAGTG 1299

Qy 461 AlaLysAlaMetLysValProAlaGluGlyTyrIleMetLysAspGlyThrProTrpPro 480
 Db 1300 GCGAAGSCCTCAAGGTTCCTCTAGAAGTTGGATCATGCAAGATGGAACACCGTGGCCA 1359

Qy 481 GlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGly 500
 Db 1360 GGGAAACAACCAAGGACCAACCCCGGTATGATCAAGTCTTCTCGGCCACAGCGGAGGA 1419

Qy 501 HisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgPro 520
 Db 1420 TTTGATGTCGAGGGCATGAGCTTCTCGGCTTGTGTACGTGTCCCTGTAAGAGCGTCTCT 1479

Qy 521 GlyPheGlnHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSerAlaVal 540
 Db 1480 GGTTTTCAACACCAAGAAAGCTGCGCCCATGAATGCTCGTTCGAGTGGCAGGGCGTA 1539

Qy 541 LeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLys 560
 Db 1540 CTCACAAATGCTCTCTTTCATGCTGAACCTTGGACTGTGATCCTATGTAACCAACAGCAAG 1599

Qy 561 AlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysValCysTyr 580
 Db 1600 GCGTGAGGAGCAATGTTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1659

Qy 581 ValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsnArgAsn 600
 Db 1660 GTTCAGTTCCTTCAAGGTTTGTATGTCATGTGACAAACGATCGTTACGCCACAGAAAC 1719

Qy 601 ThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProValTyrVal 620
 Db 1720 ACAGTCTCTTCTGATATCAATATGAAGGCTAGATGGAATCCAAAGTCCAGGTCAGTTTACGTT 1779

Qy 621 GlyThrCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProProLysGlyPro 640
 Db 1780 GGTACTGTGTGTGTTTTTCAACGACAAAGCTCTGTATGTTGTTATGAACCCCAAGGGTCTCT 1839

Qy 641 LysArgProLysMetValThrCysAspCysCysProCysPheGly-----ArgLysLys 658
 Db 1840 AAACGTCACAAAGATGATAGCTGTGTTGTTGCTTCTGCTTCTGCGGCCCGGAGAGAAT 1899

Qy 659 ArgLysHisAlaLysAspGlyLeuProGluGlyThrAlaAspMet---GlyValAspSer 677

Db 1900 AGAAAAATTTCCAAAGAAATGACATGAATGGTGCAGTACGACCCCTTGGAGGAGCAGAGGT 1959
QY 678 ABPLySGluMetLeuMetSerHisMetAsnPhelGluLysArgPheGlyGlnSerAlaA1a 697
Db 1960 GATAAGAGCAATTTGATGTTTGAATGAACCTTTGAGAAACATTTGGGCATCATCCATC 2019
QY 698 PheValThrSerThrLeuMetGluGluGlyGlyValProProSerSerSerProAlaA1a 717
Db 2020 TTTGTAACCTCAACTTTGATGGAAGAAGGTGGTGTCTCCGTCATCAAGTCCTGCAGTG 2079
QY 718 LeuLeuLysGluAlaLeuHisValIleSerCysGlyTyrGluAspLysThrAspTTPGly 737
Db 2080 CTCCTTTAAAGAGCAATCCATGTCATGAAGTCGGTTATGAAGACAGACTGAGTGGGA 2139
QY 738 LeuGluLeuGlyTyrIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMet 757
Db 2140 ACTGAGCTGGTGGATCTATGGCTCTATCACAGAGATATTTGACGGATTCAGATG 2199
QY 758 HisCysArgGlyTyrArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySer 777
Db 2200 CATTCGGTGGATGGAGGTCTATTACTGCATGCCCTAAGAGGCGCTCAATCAAGGTTCA 2259
QY 778 AlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerVal 797
Db 2260 GCTCCTATTAACTATCAGACAGTTAAACAGGTTTGGCTGGGCACCTTGGATCGGT 2319
QY 798 GluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLys 817
Db 2320 GAGATATTTTACGCGGCACAGTCTCTCTGTATGGCTACAAAGAGGCAACTCAAG 2379
QY 818 TrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeu 837
Db 2380 TGGCTTGGCGTTTGTCTATGCAACACACAACTCTACCCCTTCACATCTATACCATT 2439
QY 838 LeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSer 857
Db 2440 CTTGCCCTACTGTATCTTCCAGCAGTCTGTCTCTTACTGACAAATTCATCGCCACCG 2499
QY 858 IleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGly 877
Db 2500 ATAAGCACATTTGCTAGTCTCTCTTCACTCACTGTTATGTCGATCATTTGAACGGA 2559
QY 878 IleLeuGluMetArgTrpSerGlyValSerIleGluGluTyrTrpArgAsnGluInPhe 897
Db 2560 ATCTTGGATAGATGGAGCGAGTAGCATTTAGCAATGGTGGAGAAACGAGCAATTC 2619
QY 898 TrpValIleGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysVal 917
Db 2620 TGGGTCATTGGAGGAATCTCAGCTCATCTCTTTGGCGTGTGTCGAAGGTCTCTCAAAATC 2679
QY 918 LeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspAsp 937
Db 2680 TTAGCAGGCATTGACACAACTTCACCGTCACATCAAGGCAACA-----GATGATGAT 2733
QY 938 GluPheAlaGluLeuTyrAlaPheLysTrpThrLeuLeuIleProProThrThrLeu 957
Db 2734 GACTTTGGAGAACTTTAGCATTTCAATGACAACTGCTGATCCTCCTCAACAACTGTC 2793
QY 958 LeuIleLeuAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAnGlyTyr 977
Db 2794 TTAATCATAAACATTTGGCGTGTGTCAGGATCTCAGATGCCATTAACATGGATAT 2853
QY 978 GlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeu 997
Db 2854 CAGTCTTGGGAGCTCTATTTTGGTAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2913
QY 998 TyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIle 1017
Db 2914 TACCCATTCCTCAAGGCTGTGATGGGTAGACAGACAGACACCAACCATTTGTTGAT 2973
QY 1018 TrpSerIleLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIle 1037

Db 2974 TGGTCAGTGTATTATGGCATCTATCTCTCTTTGCTTTGGGTAAAGATGATCCTTTTGTG 3033
QY 1038 ValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3034 CTCACAGCAAGGACCTGACACTTCCAAGTGTGGCATCACTGC 3078
RESULT 10
AY639654 3911 bp mRNA linear PLN 29-JUN-2004
LOCUS Pinus radiata cellulose synthase catalytic subunit (CesA1) mRNA,
DEFINITION complete cds.
ACCESSION AY639654
VERSION AY639654.1 GI:49182339
KEYWORDS Pinus radiata (Monterey pine)
SOURCE ORGANISM
Pinus radiata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus;
REFERENCE 1 (bases 1 to 3911)
AUTHORS Klein,A.S., Tibbits,J., Steven,R. and Anthony,B.
TITLE Cellulose synthase genes in Conifers: what we know and what we need
JOURNAL to learn
(in) Havaehi,T. (Ed.);
PLANT CELL WALLS;
2 (2005) in press
REFERENCE 2 (bases 1 to 3911)
AUTHORS Klein,A.S., Tibbits,J., Steven,R. and Anthony,B.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2004) Biochemistry and Molecular Biology,
University of New Hampshire, 46 College Road, Durham, NH 03824, USA
FEATURES
Location/Qualifiers
1..3911
/organism="Pinus radiata"
/mol_type="mRNA"
/db_xref="taxon:3347"
gene 1..3911
/gene="CesA1"
/note="PrCesA1"
5' UTR 1..168
/gene="CesA1"
CDS 169..3423
/gene="CesA1"
/EC_number="2.4.1.12"
/note="glycosyltransferase"
/codon_start=1
/product="cellulose synthase catalytic subunit"
/protein_id="AAT57672.1"
/db_xref="GI:49182340"
/translation="MEASAGLVAGSHNRNEFVVIHGHEGPKPILNTLSGHVCQICGSDV
GLNTDGEUFVACNECGFPVCRPCYEYERREGNSCPQCNTRYKQKQSPRVGGDDDEE
DYDDIEHFNVEAQNRQQTTEANLHGMSYGRGPDNDSQIAHNPELPPOIPLVAN
LKDQPTFVKERRAMKREYEFKVINRLVAKASKVPEKGTMDQGTWPQGNTRDH
PGMIQVFLHSGGLDTGEGNLPRLVVSREKRPQGHKKAGAMNALVRVSAVLTNAP
FMLNDCHVINNSKAIREGMCFMMDPOVRKVCYVQRPQFDGIDRNDRYANRNTVF
FINMKGLDGIQPVYVGTGCMFRQALYGYGPKPKPKPMVTCDCIPLCCGPKKSP
KNSSKSGAPAPAYNLGDIIEEGVEYDDEALLMSQLDPEKKFQSSAFVOSTLME
NGQVPTANPAELKEAHIVISCGYEDKTEWKELGWIYGSVTEIDLTFKMHTRGWR
STYCMKPAAPAEKSAIPINSLDRNLQVRLWALGSVEIFMSRHCFIWTYGGGLKWLRF
AYINTIVPTSLPIIAYCTLPASVLLTKFVIPQISTPASLFFIATLFIATGLILE
MEWSGVSIEEWRNQFVIGVSAHFAVIGQLKVLAGIDTNFTVTAKASDDGFG
ELYAKPTWTLIIPPTILLVNLGVVGVADAINNGPQSGPLGLKLFAPFWVILHY
PFLKGLMGKQKRPITIVIWSILASVLSFVWRIDPFLSKVGPDKQCGINC"
3424..3879
/gene="CesA1"
3' UTR

ORIGIN
Alignment Scores: 0 Length: 3911
Pred. No.:

Score: 4483.00 Matches: 841
Percent Similarity: 84.56% Conservative: 90
Best Local Similarity: 76.39% Mismatches: 104
Query Match: 79.37% Indels: 66
DB: 8 Gaps: 16

US-10-627-132-30 (1-1052) x AY639654 (1-3911)

Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
Db 169 ATGGAAGCAGCGCGCGTGGTGGCGGTTCTCATACAGAAACGAGTTCGTGTCATC 228
Qy 21 ArgGlyHisGluAspProLeuArgAlaLeuSerGlyGlnValCysGluIleCys 40
Db 229 CATGACATGAGGGCCGACGCTTGAACAGTTGAGTGCCACGCTGCCAGATTGT 288
Qy 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
Db 289 GCGAGGAGCGTGGGCTTAACACAGACGCGGAGCTGTTGCTGTGTAATGAGTGGCGG 348
Qy 61 PheProValCysArgProCysTyrrGluArgGluGlyThrGlnAsnCysPro 80
Db 349 TTCTCTGTCTGCGCGCGTGTATGAGTACGAGACGAGAGGAATCATCGTGGCCG 408
Qy 81 GlnCysLysThrArgTyrrLysArgLeuLysGlySerProArgValAlaGlyAspAsp 100
Db 409 CAGTGCAATATCTGTTTCAACGCGTCAAAAAGGGAGTCCACGGGTGGAGGTGACGATGAT 468
Qy 101 GluGluAspIleAspAspLeuHisGluPheAsnIleAspAspGluAsnGlnArg 120
Db 469 GAAGAGAGCGTTGATGACATAGACATGAATTAATGTGGAGGCTCAGCAAGAAACAGG 528
Qy 121 GlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMet 140
Db 529 CAG-----CAGATCACCGAGGCGATGCTCCACGCGCGCATG 564
Qy 141 SerTyrrGlyArgGlyProAspAspGlyAspGlyAsn-----AsnThrProGlnIlePro 158
Db 565 AGCTATGGCCGAGGTCCGCGACGAGCAAAATTCGAGATGTCATATAATCCAGACCTTCCT 624
Qy 159 -----ProIleThrGlySerArgSerValProValSerGlyGluPheProIle 175
Db 625 CCGCAGATTCTGTACTTGCAACCGGCCACTCGGT-----GTGAGTGGGAGATTCCA--- 678
Qy 176 ThrAsnGlyTyrrGlyHisGlyGluVal-----SerSerSerLeuHisLysArgIleHis 193
Db 679 ACGTCATACTACGACAGCAACCAATGTCTGCCAACCCCTGCAATGCTGAAGCGTGTGCAT 738
Qy 194 ProTyrrProValSerGluProGlySerAlaLys-----TrpAspGluLysLysGlu--- 210
Db 739 -----CCAGCTCCGAGCCGGGAGTGGAGGATCATCATGATGCCAACAGGGATATT 792
Qy 211 -----ValSerTrpLysGluArgMetAspAspTrpLysSerLys 223
Db 793 GGTTCATTATGCTTTGGGAACGTGTCTTGGAGAGGAGGAGTGATGTTATTAATTCGAAG 852
Qy 224 GlnGlyIleLeuGly-----GlyGly--- 230
Db 853 GAAAAACAAATCAGGCCAGTTGGATATGACGGAAGGGAGATATCAATATAATGGGGGTTT 912
Qy 231 -----AlaAspProGluAsp---MetAspAlaAspValAlaLeuAsnAspGluAlaArg 247
Db 913 GCACCAATAGCCTGAAGATTATATTGATCCCATATGCCAATGACCGATGAAGCAAGG 972
Qy 248 GlnProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrrArgMetVal 267
Db 973 CAGCCACTGTCCGAAAAGTCCCAATTCCTTCAAGCAAAATAATTCATACCGAATGGTC 1032
Qy 268 IleValValArgLeuValValAlaPhePheLeuArgTyrrArgIleLeuHisProVal 287
Db 1033 ATTGTAATTCGACTGATAGTCTGGTATTTTCTCCGCTATCTCTCCTGTAATCCAGTG 1092
Qy 288 ProAspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSer 307

Db 1093 AAGAAATGCATATGGGCTCTGGGCACCTTCTATCGTTTGTGAAATCTGGTTGGCTTGTCA 1152
Qy 308 TrpIleLeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrrLeuAspArg 327
Db 1153 TGGATTCCTGATCAGTTTCCCAAGTGGTGCCTATCATAGTGGTGAACGATCTTGTATCGA 1212
Qy 328 LeuSerLeuArgTyrrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPhe 347
Db 1213 CTGTGATTAAGGTACGAAACGAGAGGCGAACCATCAATGCTGTGCACCTGTGTGACCTCTTT 1272
Qy 348 ValSerThrValAspProLeuLysGluProLeuValThrAlaAsnThrValLeuSer 367
Db 1273 GTGAGTACTGTAGATCCACTGAAGAGCCCTCTTGGTTTACTGCTCAATACAGTATTATCA 1332
Qy 368 IleLeuAlaValAspTyrrProValAspLysValSerCysTyrrValSerAspAspGlyAla 387
Db 1333 ATCTCTTCAGTGTAGTACTACCTGTAGACAATGTGTCTCTGTATGTCCTGTGATGACGGAG 1392
Qy 388 SerMetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValPro 407
Db 1393 TCGATGCTTACTTTTGAATCTCTCTGTAGACCTCAGAAATTTGCGCAGAAAATGGGTACCA 1452
Qy 408 PheCysLysLysPheGlyIleGluProArgAlaProGluPheTyrrPheSerLeuLysVal 427
Db 1453 TTCTGCAAGAAATTCGACATGTAGCCCTCGCGCTCCCGAAATCTATTCTCTCAGAAAAT 1512
Qy 428 AspTyrrLeuLysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArg 447
Db 1513 GACTATCTGAAGGACAAATTTCAACCACCTTTGTCAAGAGCGCCGCGGCAATGAGAGA 1572
Qy 448 GluTyrrGluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValPro 467
Db 1573 GAATATGAAGAAATTCAGGTGCGCATCAATCGGTGTTGTTGCAAAAGGCTCTAAAGTGCC 1632
Qy 468 AlaGlyLysTrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHis 487
Db 1633 AAGGAAGATGGACAAATGCAAGACGGTACCGCTTGGCTGGTAAATAATACCCGTGACCAT 1692
Qy 488 ProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGlu 507
Db 1693 CCTGGTATGATCCAGTGTCTTGGGTACAGTGGCGGCTCGCATACAGAAAGCAATGAG 1752
Qy 508 LeuProArgLeuValTyrrValSerArgGluLysArgProGlyPheGlnHisLysLys 527
Db 1753 CTTCCTCGGCTAGTATATGTTCTCGTGAAGAGACCTGGTTTCCAGCATCACAGAAG 1812
Qy 528 AlaGlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMet 547
Db 1813 GCGCGTGCCATGAATGCTTTGGTTGCGGTTTCTGCTGTGCTCACCAGTGTCCATTATG 1872
Qy 548 LeuAsnLeuAspCysAspHisTyrrIleAsnAsnSerLysAlaIleArgGluAlaMetCys 567
Db 1873 CTGAATCTGGATGTGTGATCACTATCAACAGCAAGCAATCAGGAAAGCATGTGC 1932
Qy 568 PheLeuMetAspProGlnValGlyArgLysValCysTyrrValGlnPheProGlnArgPhe 587
Db 1933 TTTATGATGATCCTCAGGTGGGAGAAAAGTCTGTATGTCTCCAAATTCCTCCTCAGAGATTC 1992
Qy 588 AspGlyIleAspValHisAspArgTyrrAlaAsnArgAsnThrValPheAspIleAsn 607
Db 1993 GATGGTATTGATCGCAATGACCGTTACGCAATCGAAACCCGATTCTTTTGATATCAAC 2052
Qy 608 MetLysGlyLeuAspGlyIleGlnGlyProValTyrrValGlyThrGlyCysValPheArg 627
Db 2053 ATGAAGGTCTGGATGGAATTCAGGGCTGTATATGTGGGAACGTGATGATGATTCAGA 2112
Qy 628 ArgGlnAlaLeuTyrrGlyTyrrAsnProProLysGlyProLysArgProLysMetValThr 647
Db 2113 AGACAAGCTCTATATGGGTATGGGCTCTCCCAAGGCCCAAAACGTCCTCAAGATGGTGACC 2172
Qy 648 CysAspCysCysProCysPheGlyArgLysArgLysHisAlaLys----- 663

```
Db 2173 TGTGATTGTCCTCCCTGTTGGGTCCTCGTAAGAAGTCTCCGAAGAAAAATAGTAGCAAG 2232
Qy 664 -----AspGlyLeuProGluGlyThrAla 671
Db 2233 AAAAGTCAGGAATCCAGCTCCCGCTCAAACTCGACCGGATCGAGGAGAGTAGAA 2292
Qy 672 AspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAenPheGluLysArg 691
Db 2293 -----GGTATGATGACGAAGAAGCATTGTCATGACCAACTAGACTTCGAGAAGAAG 2346
Qy 692 PheGlyClnSerAlaAlaPheValThrSerThrLeuMetGluGlyGlyValProPro 711
Db 2347 TTGGCCAGGTCCTTCAGCTTTTGTTCATCCACTCTGATGGAGAATGGTGTTCGCGCA 2406
Qy 712 SerSerProAlaAlaLeuLeuLysGluAlaLeuHisValIleSerCysGlyTyTrGlu 731
Db 2407 ACAGCAATCCAGTGAATTTGAGGAGGCTATTATGTCATCATGCTGTGATATGAA 2466
Qy 732 AspLysThrAspTrpGlyLeuGluLeuGlyTrpIleTyTrGlySerIleThrGluAspIle 751
Db 2467 GACAAAACGGAATGGGGAAGAGCTTGGATGATCTATGGATCAGTCACAGAGGACATT 2526
Qy 752 LeuThrGlyPheLysMetHisCysArgGlyTrpArgSerValTyTrCysMetProLysArg 771
Db 2527 CTGACTGGATTCAAGATGCACATCGAGGCTGGCGTCCATTTACTGTATGCCCAACGA 2586
Qy 772 AlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArg 791
Db 2587 GCAGCATTCAGAGGCTGCTCCATCAATCTATCAGACCGTTTGAACAGGTTTGGCT 2646
Qy 792 TrpAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyTrGlyTy 811
Db 2647 TGGGCTTTGGGATCAGTAGAAATTTTCATGAGCAGACATTCGCCCAATCTGCTATGCTAT 2706
Qy 812 LysAsnGlyAsnLeuLysTrpLeuGluArgPheAlaTyTrIleAsnThrThrIleTyTrPro 831
Db 2707 ---GGGGAGGCTGGAATGGCTTGAAGATTTGGCTATATCAACCACTATGCTATCCA 2763
Qy 832 PheThrSerLeuProLeuLeuAlaTyTrCysThrLeuProAlaValCysLeuLeuThrGly 851
Db 2764 TTCACCTCTCTCCACTCATGCTATTGGACACTTCCAGCCGTCAGTTTGTCTCATGGC 2823
Qy 852 LysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheLeAlaLeuPheMet 871
Db 2824 AAATTTGTGATCCCTCAGATCAGTACTTTTGCAGTCTATTTTATAGCTCTTTTCATC 2883
Qy 872 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluTyTrp 891
Db 2884 TCAATTTTGGCCACTGGTATTTCTGGAAATGAGGTGGAGTGGAGTGAAGCAATGG 2943
Qy 892 TrpArgAsnGluGlnPheTrpValIleGlyValSerAlaHisLeuPheAlaValVal 911
Db 2944 TGGCGAATGAACAGTTCTGGGTTATTTGGAGGGTTTCTGCACATTTTTCAGTTATT 3003
Qy 912 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 931
Db 3004 CAAGGTCGTCTCAAGGTACTTGGCAGGCATTGTATACAAATTTTCAGTCACTGCCAAGCA 3063
Qy 932 ThrGlyAspGluAspGluPheAlaGluLeuTyTrAlaPheLysTrpThrThrLeuLeu 951
Db 3064 TCA-----GATGACGGTGGTATTGGGGAACTGTATGCATTCAAATGGACCACTCCCT 3117
Qy 952 IleProProThrThrLeuLeuIleAsnValIleGlyValValAlaGlyIleSerAsp 971
Db 3118 ATTCTCTTACACCTGCTTGTTCATCAACTTGTGGGGTGGTGTGTGGCGTAGCAGAT 3177
Qy 972 AlaIleAsnAenGlyTyTrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPhe 991
Db 3178 GCAATCAACAATGGATTTGATCATGTCATCAACTTGTGGGGTAAAGCTTTTCTTTGCATTC 3237
Qy 992 TrpValIleValHisLeuTyTrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1011
Db 3238 TGGGTCAATTGTGACCTGTATCTTTCTCAAGGGTCTCATGGGCGAGCGAAGACCGAACA 3297
```

```
Qy 1012 ProThrValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpVal 1031
Db 3298 CCCACCATCGTGTATTGGTCAATTCCTGCTGGCATCTGTTTCTCTCTCTCTCTGGTA 3357
Qy 1032 ArgIleAspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIleAsn 1051
Db 3358 AGAATTTGATCCTTCTTCTGAGTAAGGTTAAAGGCCAGATACTAAACAATGTGCATCAAC 3417
Qy 1052 Cys 1052
Db 3418 TGC 3420

RESULT 11
BD236020 3851 bp DNA linear PAT 17-JUL-2003
LOCUS Materials and method for modification of plant cell wall
DEFINITION Polysaccharides.
ACCESSION BD236020
VERSION 1 GI:33045790
KEYWORDS JP 2002527056-A/42.
SOURCE Pinus radiata (Monterey pine)
ORGANISM Pinus radiata
REFERENCE 1 (bases 1 to 3851)
AUTHORS Bloksberg,L.N.
TITLE Materials and method for modification of plant cell wall
JOURNAL Patent: JP 2002527056-A 42 27-AUG-2002;
GENESIS RESEARCH AND DEVELOPMENT CORP LTD, FLETCHER CHALLENGE
FORESTS LTD
COMMENT OS Pinus radiata (radiata pine)
FN JP 2002527056-A/42
PD 27-AUG-2002
PF 08-OCT-1999 JP 2000575985
PR 13-OCT-1998 US 09/170862,11-AUG-1999 US 60/148426 PI
LEONARD NATHAN BLOKSBERG
PC C12N15/09,A01H5/00,C12N5/10,C12N9/00,C12N15/00,C12N5/
PC 00
CC Materials and method for modification of plant cell wall CC
FH Key Location/Qualifiers
FT source 1..3851
FT /organism='Pinus radiata (radiata pine)'.
FEATURES
source 1..3851
/organism='Pinus radiata'
/mol_type='genomic DNA'
/db_xref='taxon:3347'
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3851
Score: 4477.00 Matches: 839
Percent Similarity: 84.56% Conservative: 92
Best Local Similarity: 76.20% Mismatches: 104
Query Match: 79.27% Indels: 66
DB: Gaps: 16
US-10-627-132-30 (1-1052) x BD236020 (1-3851)
Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
Db 153 ATGGAACCGACCGCGCTTGGTTGCCGTTCTCATACAGAAACGAGTTCGTGGTCATC 212
Qy 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyClnValCysGluIleCys 40
Db 213 CATGGACATGAGAGCGGAAGCCTTTGAACACAGCTTGAGTGGCCACGCTGCCAGATTGT 272
Qy 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
Db 273 GCGGAGGAGCTGGGCTTAAACACAGACGGGAGCTGTTGTTGCCCTGTATGATGATGCGGG 332
```

```
QY 61 PheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsnCysPro 80
Db 333 TTTCTGTGTGTCGGCGCTGTATGATGACGAGACGAGAGAAATCAGTCGTGCGCG 392
QY 81 GlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAspAsp 100
Db 393 CAGTCAATACCTCGTTTACAGCGTCAAAAGGAGTCCACGGGTGGAAGGTGACCATGAT 452
QY 101 GluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnArg 120
Db 453 GAAGAAGAGCGTTGATGATAGACATAGAAATTAATGTGGAGACTCAGCAAAAGAACAGG 512
QY 121 GlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMet 140
Db 513 CAG-----CAGATCACCGAGCGCATCTCCACGAGCGCATG 548
QY 141 SerTyrGlyArgGlyProAspAspGlyAspGlyAsn-----AsnThrProGlnIlePro 158
Db 549 AGCTATGCGGAGGTCCGACGACGAAATTCGAGATGCTCATAAATCCAGAGCTTCCT 608
QY 159 -----ProIleIleThrGlySerArgSerValProValSerGlyGluPheProIle 175
Db 609 CGCAGATTCCTGTACTTGCAACCGCCACTCGGTT---GTGAGTGGGAGATTCCA--- 662
QY 176 ThrAsnGlyTyrGlyHisGlyGluVal-----SerSerSerLeuHisLysArgIleHis 193
Db 663 ACGTCATATACCGCAGACAAACCAATGTCTTGCAACCCCTGCAATCTGAAGCGTGTGCAT 722
QY 194 ProTyrProValSerGluProGlySerAlaLys-----TrpAspGluLysGlu---- 210
Db 723 -----CCAAGCTCCGAGCGGGAGTGGAGGATCATCATGATCCAAACAGGGATATT 776
QY 211 -----ValSerTrpLysGluArgMetAspAspTrpLysSerLys 223
Db 777 GGTTCCTATGCTTGGGNAAGTGTCTTGGAAGGAGCGGCGATGTTATATAATCGAAG 836
QY 224 GlnGlyIleLeuGly-----GlyGly----- 230
Db 837 GAAAAAATAATCAGGCGAGTGGATATGACGGAAGGGAGATATCAATATAATGGGGGTTT 896
QY 231 -----AlaAspProGluAsp---MetAspAlaAspValAlaLeuAsnAspGluAlaArg 247
Db 897 GCACCAATATGACCTTGAGATTTATTTGATCCCGATATGCCAATACCGATGAGCAAGG 956
QY 248 GlnProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetVal 267
Db 957 CAGCCACTGTCGCGAAAGTGCATTCCTTCAAGCAAAATAAATCCATCCGATGGTC 1016
QY 268 IleValValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProVal 287
Db 1017 ATTTGTAATTCGACTGATAGTGTGGGTATTTTCTCCGCTATCTGCTCTGTAATCCAGTG 1076
QY 288 ProAspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSer 307
Db 1077 AAGATGATATATGGGCTCTGGGCCACTTCTATCGTTTGTGAAATCTGGTTGGCTTGTCAT 1136
QY 308 TrpIleLeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArg 327
Db 1137 TGGATTCCTTGATCAGTTCCTCCAAAGTGTGCTATCATCAGTCGTGAAACGTAICTTGATCGA 1196
QY 328 LeuSerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPhe 347
Db 1197 CTGTCAATTAAGGTAGCAACAGAGCGCAACCATCAATGCTTGACCTGTGGACCTCTTT 1256
QY 348 ValSerThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSer 367
Db 1257 GTGAGTACTGTAGATCCACTGAAGGAGGCGCTCTTTGGTTACTGCGCAATACAGATATATCA 1316
QY 368 IleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAla 387
Db 1317 ATCCCTTTGATGATAGACTACCTGTAGCAATGTGCTCTGTTATGCTCTGATGACGAGCG 1376
QY 388 SerMetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValPro 407
```

```
Db 1377 TCGATGCTTACTTTTGAATCTCTCTGAGACCTCAGAAATTTGCCAGAAATTTGGGTACCA 1436
QY 408 PheCysLysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysVal 427
Db 1437 TTTTCGCAAGAAATTCGACATTTGAGCTTCGCGCTCCGNAATCTATTTCTCTCAGNAATTT 1496
QY 428 AspTyrLeuLysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArg 447
Db 1497 GACTATCTGAAGGACAAATTTCAACCCACCTTTGTCAAAAGAGCGCGTGCATGAAGAGA 1556
QY 448 GluTyrGluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValPro 467
Db 1557 GAATATGAAGAAATTTCAAGGTGCGCATCAATCGGTGGTGTGTCAAAGGCCCTCTAAAGTGCC 1616
QY 468 AlaGluGlyTrpIleMetLysAspGlyThrProTyrProGlyLysAsnThrArgAspHis 487
Db 1617 AAGNAGGATGACCAATGCAAGACGCTAGCGCTTGGCTGGTATATATACCGTGACCAT 1676
QY 488 ProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGlu 507
Db 1677 CTTGGTATGATCAAGTGTCTTGGGTCAACAGTGGCGCTCGATACAGAAAGGCAATGAG 1736
QY 508 LeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLys 527
Db 1737 CTTCTCGGCTAGTATATGTTCTCTGAGAGAGACCTGGTTTCCAGCATCACAAAGAG 1796
QY 528 AlaGlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMet 547
Db 1797 GCCCGTGCATGAATGCTTTGGTTCGGGTTTCTGCTGTGCTCCCAATGCTCCATTTATG 1856
QY 548 LeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCys 567
Db 1857 CTGAATCTGGATTTGTGATCACTACATTAACAATAGCAAGCAATCAGCAAGGCGCATGTGC 1916
QY 568 PheLeuMetAspProGlnValGlyValGlyValCysTyrValGlnPheProGlnArgPhe 587
Db 1917 TTTATGATGGATCTCTCAGTTTGGGAGAAAGTCTGTATGTCCTCAATTCCTCCAGAGATT 1976
QY 588 AspGlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsn 607
Db 1977 GATGGTATTGATCGCAATGACGCTTACGCCAATCGAAACCGTATTTCTTTGATATCAAC 2036
QY 608 MetLysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArg 627
Db 2037 ATGAAAGCTCTGGATGGAATTCAGAGGCTCTATATGTTGGAACTGGATGATGTTTCA 2096
QY 628 ArgGlnAlaLeuTyrGlyTyrAsnProLysGlyProLysArgProLysMetValThr 647
Db 2097 AGACAAAGCTCTATATGGGTATGGGCTCTCCAAAGGCGCCAAACCGTCCCAAGATGGTGACC 2156
QY 648 CysAspCysCysProCysPheGlyArgLysLysArgLysHisAlaLys----- 663
Db 2157 TGTGATGTCTCCCTTGTGTGGTCTCTCGTAAGAAGTCTCCGAGAGAAATAATAGTAGCAAG 2216
QY 664 -----AspGlyLeuProGluGlyThrAla 671
Db 2217 AAAAGTGCAGGAATCCACAGCTCCCGCTTACAATCTGGACGGATCGAGAGGAGTAGAA 2276
QY 672 AspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArg 691
Db 2277 -----GGTTATGATGAGAAAGAGCATTTGTGTATGAGCCCACTAGACTTCGAGAGAG 2330
QY 692 PheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGluGlyValPro 711
Db 2331 TTTGGCCAGTCTTCAGCTTTTGTTCATCCACTCTGATGGAGAAATGGTGGTGTCCGCA 2390
QY 712 SerSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGlu 731
Db 2391 ACAGCAAAATCCAGCTGAATTTGTAAGGAGGCTATTCATGTCATCAGCTGTGGATATGAA 2450
QY 732 AspLysThrAspTrpGlyLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIle 751
```

```
Db 2451 GACAAACCGAATCGGGAAAGAGCTTGGATGGATCTATGGATCAGTCACAGAGACATT 2510
Qy 752 LeuThrGlyPheLeuMetHisCysArgGlyTrpArgSerValTyrCysMetProLeuArg 771
Db 2511 CTGACTGGATTCAAGATGCACACTCGAGGCTGGGGTCCATTACTGTATGCCCAACGA 2570
Qy 772 AlalaPheLeuGlySerAlaProIleAsnLeuSerAspArgLeuAenGlnValLeuArg 791
Db 2571 GCAGCATTCAAAGGCTCGCTCCAAATCAATCATCAGACCGTTTGAACCGAGTGTGGT 2630
Qy 792 TrpAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyr 811
Db 2631 TGGGCTTGGGATCAGTAGAAATTTTCATGACGACACATTCGCCAATCTGGTATGGGTAT 2690
Qy 812 LysAsnGlyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 831
Db 2691 ---GGGGAGGCTCGAATGGCTTGAAGATTTCCTATATCAACACCATTCGTATCCA 2747
Qy 832 PheThrSerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGly 851
Db 2748 TTCACCTCTCTTCCACTCAATGCTATTCACACTTCACAGCGTCAGTTGCTCACTGGC 2807
Qy 852 LysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMet 871
Db 2808 AATTTGTGATCCCTCAGATCAGTACTTTTGCAGAGCTTATTTTAAATAGCTCTTTTCATC 2867
Qy 872 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTrp 891
Db 2868 TCATTTTGGCACTGGTATTCTGAATGAGGTGGAGTGAGCATGGAGCATGG 2927
Qy 892 TrpArgAsnGluGlnPheTrpValIleGlyValSerAlaHisLeuPheAlaValVal 911
Db 2928 TGGCGAAATGAACAGTCTCGGTATTGGAGGGTTCCTGCACATTTTTCGACGTTATT 2987
Qy 912 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 931
Db 2988 CAAGGCTGCTCAAGGTACTGGCAGGCATTGATACAAATTTTCAGTCACCTGCCAAGGCA 3047
Qy 932 ThrGlyAspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeu 951
Db 3048 TCA-----GATGACGGTGAGTTGGGGNACTGTATGCATTCAAATGGACCACTCTC 3101
Qy 952 IleProThrThrLeuLeuIleIleAsnValIleGlyValValAlaGlyIleSerAsp 971
Db 3102 ATTCCTCTCTACACCTCTGTCTCATCAACCTTGTGGGGTGGTGTGGTGGTAGCAGAT 3161
Qy 972 AlaIleAsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPhe 991
Db 3162 GCATCAACAAATGGATTTTCAGTCATGGGGTCTCTCTTGGGTAAAGCTTTTCTTTGGCATC 3221
Qy 992 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1011
Db 3222 TGGGTCAATTGGCACCTGTATCTCTTCTCAAGGGTCTCATGGGCGAGACCGACCA 3281
Qy 1012 ProThrValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpVal 1031
Db 3282 CCCACCATCGTGGTATTGTGTCATTCGTGGCATCTGTTTCTCTCTTTCTGGGTA 3341
Qy 1032 ArgIleAspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIleAsn 1051
Db 3342 AGAATTGATCTTTCTTGGTAGGTTAAAGGCCCAAGATATAAACAATGTGGCATCAAC 3401
Qy 1052 Cys 1052
Db 3402 TGC 3404
LOCUS AC140546 106962 bp DNA linear PLN 09-MAR-2004
DEFINITION Medicago truncatula clone mth2-35119, complete sequence.
ACCESSION AC140546
VERSION AC140546.10 GI:45268058
KEYWORDS HTG.
```

SOURCE
ORGANISMMedicago truncatula (barrel medic)
Medicago truncatulaEukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.REFERENCE
AUTHORS

1 (bases 1 to 106962)

Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.

Medicago truncatula BAC Clone mth2-35119

TITLE
JOURNAL

Unpublished

REFERENCE
AUTHORS

2 (bases 1 to 106962)

Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.

Direct Submission

Submitted (25-FEB-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USAREFERENCE
AUTHORS

3 (bases 1 to 106962)

Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.

Direct Submission

Submitted (30-SEP-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USAREFERENCE
AUTHORS

4 (bases 1 to 106962)

Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.

Direct Submission

Submitted (16-OCT-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USAREFERENCE
AUTHORS

5 (bases 1 to 106962)

Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.

Direct Submission

Submitted (09-MAR-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

COMMENT

On Mar 9, 2004 this sequence version replaced gi:37693616.

----- Genome Center

Center: Department Of Chemistry And Biochemistry

The University Of Oklahoma

Center code:UOKNOR

FEATURES
source

Location/Qualifiers

1..106962

/organism="Medicago truncatula"

/mol_type="genomic DNA"

/db_xref="taxon:3880"

/clone="mth2-35119"

/clone_lib="Medicago truncatula BAC library H2"

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	106962
Score:	4247.50	Matches:	863
Percent Similarity:	60.86%	Conservative:	90
Best Local Similarity:	55.11%	Mismatches:	83
Query Match:	75.20%	Indels:	534
DB:	8	Gaps:	21

US-10-627-132-30 (1-1052) x AC140546 (1-106962)

Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
|||||
Db 26794 ATGGAAGCCAGCCGCCGAGCTAGTGGTGTCTCAACACCGGAACGAGCTTGTGTCATT 26853
|||||Qy 21 ArgGlyHisGluAspPro----- 26
|||||Db 26854 CATGGCCATGAGAGAGGT-ATAATAAACTTATAACCAAGTGGTGGTAGAGATATTT 26912
|||||

Qy 27 -----LysProLeu 29

Db	26913	CTTTGTTGTTTATGTTGATAGTTTGTGTTGTTGTTTGTGTCAGCAGCAAGCCATG	26972
QY	30	ArgAlaLeuSerGlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAsp	49
Db	26973	AAAGAACTTGGACGGTCAAGTGTGCGAGATATGTTGGTATGATGATGTTGGACTTACTGTGGAT	27032
QY	50	GlyAspLeuPheValAlaCysAsnGluCysGlyPheProValCysArgProCysTyrGlu	69
Db	27033	GGAGATTTGTTGTGGCATGTAAACGAGTGTGTTTTCGGTGTGCGGCTTGTCTATGAG	27092
QY	70	TyrGluArgGlyThrGlnAsnCysProGlnCysLysThrArgTyrLysArgLeu	89
Db	27093	TATGAAGAAGGGAGGGAGACAATTGTCTCAGTCAAGACCAAGATACAGAGGTCTC	27152
QY	90	-Lys	90
Db	27153	AAAGGTAGAATTATATCTGAACAAGTTATTTTAGCCCTTGTGTTATCATGCAATAAATT	27212
QY	90		90
Db	27213	GTAAACATGAACCTTACATGATACAGTTTGTAGATGACTGTGATATTAATTTGATCGAT	27272
QY	91		100
Db	27273	TAACCTAAATGATATACTAATATGATGTTTAGGGACCTCTGGGTTGAGGAGACATGA	27332
QY	100	pGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnAr	120
Db	27333	TCAGGAGGATGTGGACGATATTGAACATGAATTCAAAATTGAAGACAA	27381
QY	120	gGlnLeuGluGlyAsnMetGlnAsn--SerGlnIleThrGluAlaMetLeuHisGlyAr	139
Db	27382		149
QY	139	gMetSerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIleProPr	159
Db	27426	GATGAGCTATGAAGAGGTCTGAAGAT-----GATGAAATGACACTTCCACG	27476
QY	159	oIleIleThrGlySerArgSer-----	166
Db	27477	TGTTATTGTGTGTGTCGCTCGGAAATGAAGATTTTCCACCATTTCTAGTCCATCTAA	27536
QY	167		170
Db	27537	ACTCGATAGAAGCTGTGAATTCATACAACTAAATTTGCTTTTATACATGTCAGGTGAG	27596
QY	170	rGlyGluPheProIle---ThrAsnGlyTyrGlyHisGlyGluValSerSerLeuHi	189
Db	27597	TGGTGAATTCCTCAATATCATCTAGTTATGGGAG---CAGATGCTATCTTCACTGCA	27653
QY	189	sLysArgIleHisProTyrProValSerGlu-----	199
Db	27654	TAAAGAGTGCATCCATATTCGCTTCATGTTCTGTAAGATGTAGTTATAGATTATC	27713
QY	199		199
Db	27714	ACTTTCTGTTTCATTTGCAAAACGGATTATGCCAAATTTGTTAAATTTCTTGTCTAAAG	27773
QY	199		199
Db	27774	TCCAAATCTTACTAGTTATACAGCCCTTTCTTTTGGCAGTGAATAAAATCAACCCCTG	27833
QY	199		199
Db	27834	AAATGCTTTTCTTAGCTTACTACACATGGTGGATATGATTATATATGTGGCATGCTGT	27893
QY	200	-----ProGlySerAlaLysTrpAspGluLysGluValSerTrpLysGluArgMe	217
Db	27894	GATAAAACAGGAAGTGCAGGTTGGATGAAAGGAGAGAGATGATCATATATAGAAAT	27953
QY	217	trpAspTrpLysSerLysGlnGlyIleLeuGlyGlyAlaAspProGluAspMetAs	237
Db	27954	GGATCACTGGAAATTTGCAGCAAGGAAATTTGGGACCTGAACCTGAT---	28010
QY	237	palAspVal	240
Db	28011	TGCAAAACATGATGCTTTTTCAGTACTAAATGCAATTTCTCTGACAAATTTAAATCTTAATGAAA	28070
QY	240		240
Db	28071	TGATAGTAATATTTGAAAGATCATAAATTAATTAATTTAGTTACTACTATPACAATGTG	28130
QY	241		258
Db	28131	AACCCATGCAGG-TCAGATGAAGCAACCACTGTCAAGAAAGGTACCAATAGCATC	28189
QY	258	rSerLysValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPhePh	278
Db	28190	TAGCAAAATCAATCCATAGATGCTGTCGACGGCTTGTATTCTTTGGCTTCTT	28249
QY	278	eLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerIl	298
Db	28250	CCTCCGATACAGACTTATGAATCCAGTACATGATGCAATGGGGCTATGGCTAACTCAAT	28309
QY	298	eIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPhePr	318
Db	28310	TATATGCAATCTGTTTGTCTATATCATGATCTTGCATCAGTTCTCCCAATGGTATCC	28369
QY	318	oIleAspArgGluThrTyrLeuAspArgLeuSerLeu-----	330
Db	28370	TATTGATAGAGAAACATACCTTTGACCGTCTTTCTCACTCAGGTAAAGGGACCTTCCCTCCTG	28429
QY	331		331
Db	28430	ATTGTTGAAATACCACTGTTTAAATATTTATATCTCTGAACAAAGAAAGTTGTTTATCAG	28489
QY	331	gTyrGluArgGluGlyGluProSerLeuSerAlaValAspLeuPheValSerThrVa	351
Db	28490	GTAAGAGCTGAAGCGCAACCAATATGCTTCTCTCTAGATGTTGTTGTAGTACTGT	28549
QY	351	lAspProLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAlaVa	371
Db	28550	GGATCCCTTGAAGAAACCTCTCTGAATACAGCCAAACACAGTTCTTCAATCTTGGCAAT	28609
QY	371	lAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLeuTh	391
Db	28610	GGACTACCCCATGTATAAGATATCATGCTACATTTCTGATCAGGAGCTTCAATGTGAC	28669
QY	391	rPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLysLy	411
Db	28670	ATTTGAGCCCTGTCGAAACCGCAGAGTTTGTAGGAAGTGGGTACCATTTTGTAGAA	28729
QY	411	sPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeuLy	431
Db	28730	ATTTTGTATGAACCTCTGTCACCAAGAGATGTACTTCTCTGAGAAATTTGACTATCTAAA	28789
QY	431	sAspLysValGlnProThrPheValGlnGluArgAlaMetLys-----	446
Db	28790	GGCAAGGTGACGCCAATTTTGTCAAGAAACGTGGTCCATGAGAGTTTGTATTAAATCT	28849
QY	446		446
Db	28850	ATTAGATTTGTTGAAAAAATGAAGTTCTGAAATCAGTTTAAAGCTTACTCACCAGCTA	28909
QY	447		460
Db	28910	TTAATTGGATGCATGGCAGAGAGATATGAAGAGTTTAAAGTTAGATCAATGCATTTGT	28969
QY	460	lAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrProTrpPr	480
Db	28970	GGCAAAAGCTCAGAAGGTTCCTGCAGGAGGTGGATATATGAGGATGGACACCATGGCC	29029
QY	480	oGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGl	500
Db	29030	AGGAACAATATCTAAGGATCATCTCTGGTATGATTCAAGTCTTTCTTGTGTACAGTGGAGG	29089

```
QY 500 yHieAspThrGluGlyAenGluLeuProArgLeuValTyrValSerArgGluLysArgPr 520
D 29090 TCATGATAGTAGGAAACACGCTCTCTCGCTGTTTATGATCTAGAGAGAAAGGCC 29149
QY 520 oGlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeu----- 534
D 29150 AGGATTTCAACACACACAAAGCGGTGCCATGAATGCTCTGCTAGAGTTACTTTGAACA 29209
QY 534 ----- 534
D 29210 AACAGATTTTGTATTTTATGATGTCATTCACCTTTCTGAAACACACTTTACTTTGACTC 29269
QY 535 ----- 535
D 29270 ATCTAAATAATTTATGTAATCTCCAGGTACGGGTCTCTGCGATGCTTCAAAATGCTCTTT 29329
QY 546 eMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMe 566
D 29330 TATGCTGAACCTTGGATTGTGATCATTAATCAATAACACAGCAAGGCTGTCCGAGAGGCCAT 29389
QY 566 tCysPheLeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnAr 586
D 29390 GTGCTTCTTGATGGACCCCAACAGGGAAGAAAGTCTGCTATGTCCAGTTTCTCTCAAG 29449
QY 586 gPheAspGlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspI 606
D 29450 ATTGATGGCATTTGATGGCATGATCGATATGCTAATAGGAACACAGATTTTCTTTGATGT 29509
QY 606 e----- 606
D 29510 -AAGTCACTGCAAGAAAGTCAACATCAACAGCATGATCTCTTTTGAAGCAATTGATGT 29568
QY 606 ----- 606
D 29569 TTTTGTAGTAGCAAGCTAATATATATATCAAAATCTCTTGTCTGTATGCGCAGAT 29628
QY 607 -AsnMetLysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPh 626
D 29629 TAACATGAAGGGTCTAGATGTTATTCAGGGTCTCTGTATGTCGGCACAGGGTGTGTTT 29688
QY 626 eArgArgGlnAlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetVa 646
D 29689 TAGAAGGCAGGGCTTTATATGATATAATCTCTTAAGGGTCCCAAGCGTCCAAAAATGGT 29748
QY 646 lThrCysAspCysCysProCysPheGlyArgLysLysArg--LysHisAlaLysAsp- 664
D 29749 AAGTTGTGATGTCGCCATGTTTGGAAAGGCGCAAGGTTAAGCATGCAATGAACGA 29808
QY 665 -----GlyLeuProGlu----- 668
D 29809 CGCAATGGAGAGCTGCGAGGCTAAG-AGGTTCTAAATTTTACAATAAGTACGCAATCTT 29867
QY 669 -----GlyThrAl 671
D 29868 GATACTAAATCTTTATACAAATGATTTTATGTTTCACTGTTCTTAATCTCTGTACATGT 29927
QY 671 asPMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysAr 691
D 29928 CCATATAGGAATGGGAAGATGACAAAGAGTTACTGATGTCCTCCAGATGAATTTTGAGAAGA 29987
QY 691 gPheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGlyGlyValProPr 711
D 29988 ATTTGGGCAGCATCATGATTTTGTGACTTTCAGTCTTGTATGGAGAGGGTGGTGTACCTCC 30047
QY 711 oSerSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGl 731
D 30048 TTCGTCAAGTCCAGCAAGCCAACTTAAAGAAAGCCATTATGTAATCAGTTGCGGATATGA 30107
QY 731 uAspLysThrAspTyrGlyLeuGlu----- 739
D 30108 AGATAAAACTGTAATGGGGGATTTGAGGTAAACCAACATTTGGCAGCCTTAGGCTCTCAATTG 30167
QY 739 ----- 739
D 30168 ATATTTTACAATAGTTGTTTCAACAATAAGAAATCAGCTCATCCATATTTTGTTCCAA 30227
QY 740 -----LeuGlyTyrIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMe 757
D 30228 ATTGCACTTGGTGGATTTATGATCTATTACAGAGGATATTTTACAGGCTTTAAGAT 30287
QY 757 tHisCysAsArgGlyTyrArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySe 777
D 30288 GCATTGCCGTGGTGGAGATCCATTACTGTATGCCAAGAGAGTAGCATTTCAAGGTAC 30347
QY 777 rAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTyrPheAlaLeuGlySerVa 797
D 30348 TGCTCTCTATCAACTTGTGCATAGACTCAACACAGGTGCTTCGTTGGGCTCTTGGCTCAT 30407
QY 797 lGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLy 817
D 30408 TGAGATCTCTTTCAGTCATCATTTGCCCTTTATGGTATGGCCCAAGGAAGGAGCTGAA 30467
QY 817 sTyrLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLe 837
D 30468 GTGGCTGGAGCGATTTGGCTATGCNAACACACACCGGTGATCCCTTCACTCCATCTCT 30527
QY 837 uLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetPro-- 856
D 30528 AGTTGCCCTACTGTATTTCTCCAGCTGTCTGCTTACTCACTGCACAAATTCATCATGCCAC 30587
QY 856 ----- 856
D 30588 GGTAACATTTTATTTCTTTTATCTCTGTTTAGAAAATGAGTTAAGACACAGCAGCTTT 30647
QY 856 ----- 856
D 30648 TGAAATTAATAATTTTTCACAAACATTAACAATATTTTCAATATATTTACCTTCAAAAGCACA 30707
QY 856 ----- 856
D 30708 AGAATGATCAGCATCTTTTATGTATGTAGTAGGAATTCAACTTTTAAATCCATAACAGC 30767
QY 856 ----- 856
D 30768 TAGAAATGCTAGGCATCATTTGTTCAAGGATGCTCAGAGAAATTCATGCAAGCCTTTTAC 30827
QY 857 -----SerIleSerThrPheAlaSerLeuPhePheI 867
D 30828 TCAATTTTATTTTAAATTTGTTTCTTTCAGATAAGCACTTTGTTAGTTGTACTTGT 30887
QY 867 eAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValSe 887
D 30888 TGCTCTCTCTCTCAATCATGCAAGGGCAATTTTGAGTTGAAATGGAGTGGAGTCAG 30947
QY 887 rIleGluGluTyrTrpArgAsnGluGlnPheTyrValIleGlyGlyValSerAlaHisLe 907
D 30948 CATTGAGGAATGGTGGAGAAATGAGCAATCTGCGTCAATTTGTTGTTATCAGCACATCT 31007
QY 907 uPheAlaValValGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrVa 927
D 31008 CTTTGTCTGTATACAAAGGCTTCTGAAAGTTCTTGTGGAATTCATGACACAACTTCACTGT 31067
QY 927 lThrSerLysAlaThrGlyAspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTr 947
D 31068 TACATCCAAGGCAACA-----GATGATGAGGAATTTGGAGAAATGTACGCCATCAAGTG 31121
QY 947 pThrThrLeuLeuLeuProProThrThrLeuLeuIleIleAsnValIleGlyValValAl 967
D 31122 GACCACCTCTCTAATCTCTCCACGACTATCTTAATAATCAATATGTTGGGGTGTGTC 31181
QY 967 aGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTyrProLeuPheGlyLysLe 987
D 31182 TGGAACTCTCAGACCCCAATCAACATGTTACCAATCATGGGACCTCTATTTTGGAATACT 31241
QY 987 uPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyAr 1007
```



```
QY 669 -----GlyThrAl 671
Db 106271 GATACTAAATCCCTTATPACAAATTTGATTTTCATGTTTCCACCTGTTCTTAATCTCTGTGCATCATGT 106330
QY 671 aAspMetGlyValaAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysAr 691
Db 106331 CCATATAGGANTGGAGATGACAAAGATTACTGATGTCCAGATGCCAGATTAATTTTGAGANGAA 106390
QY 691 gPheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGluGlyValProPr 711
Db 106391 ATTGGGCAGTCATCGCATTTTGTGACTTCAGTCTTTGATGGAAGAGGGTGTACTCTCC 106450
QY 711 oSerSerProAlaAlaLeuLeuLysGluAlaLeuHisValIleSerCysGlyTyrGl 731
Db 106451 TTTCGCAAGTCCAGCAGCAACTTAAGAAGCCATTCATGAATCAGTTGCGGATATGA 106510
QY 731 uAspLysThrAspTyrGlyLeuGlu ----- 739
Db 106511 AGATAAACTGAATGGGGGATTTGAGGTAAACCAACATTTGGCAGCCTAGGCTCTCAATG 106570
QY 739 ----- 739
Db 106571 ATATTTTACAATAGTTGTTTACAAATAAGAAATCAGCTCATCATTTATTTTGTTCCTCAA 106630
QY 740 -----LeuGlyTyrIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMe 757
Db 106631 ATTGCAGCTTGGTGGATTTATGGATCTATTACAGAGGATATTTTAAACAGGCTTTAAGAT 106690
QY 757 thiCysArGlyTyrArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySe 777
Db 106691 GCATTGCCGTGGTGGAGATTCATTTACTGTATGTCAAAGAGAGTAGCATTTCAAGGGTAC 106750
QY 777 rAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerVa 797
Db 106751 TGCTCTCATCACTTGTGAGATAGACTCAACCCAGGTGCTTCGTGGGCTCTTGGCTCCAT 106810
QY 797 lGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLys 817
Db 106811 TGAGATCTCTTCTTCAGTCATCATTTGCCCTTTATGGTATGCCACAAAGGAAGGAGCTGAA 106870
QY 817 sTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLe 837
Db 106871 GTGGCTGGAGCGATTGGCTATGCAACACACACCGGTACCCCTTCACCTCCATCCCTCT 106930
QY 837 uLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetPro-- 856
Db 106931 AGTTGCTACTGTGTTCTTCCAGCTGTCTGCTTACTCACTGACAAATTCATCATGCCACC 106990
QY 856 ----- 856
Db 106991 GGTAAACATTTTATTCTTTTATCTTCTGTTAGAAAAATGAGTTAAGACACGACGAGCTTT 107050
QY 856 ----- 856
Db 107051 TGAATATTATAATTTTGTGACAAACATTAACAATTTTATATTTACCTTCAAAAGCACA 107110
QY 856 ----- 856
Db 107111 AGAATCATCAGCATCTTTTATGTATGTATAGTAGTAATCAACTTTTAAATCCATACAGC 107170
QY 856 ----- 856
Db 107171 TAGAAATGCTAGGCATCATTTGTTCAAGATGCTCAGAGAATTCATGCCAAGCCTTTTATC 107230
QY 857 -----SerIleSerThrPheAlaSerLeuPhePheIle 867
Db 107231 TCAATTTTTTTTTTAATTTGTTTCATTTCTTCAGATAAGCACTTTTGTAGTTGTACTTTGT 107290
QY 867 eAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValSe 887
Db 107291 TGCTCTCTCTCTTCAATCATCGCAACGGGCATTTCTTGATTTGAATGAGTGGAGTCTAG 107350
```

```
QY 887 rIleGluGluTyrTrpTrpArgAsnGluGlnPheTrpValIleGlyGlyValSerAlaHisLe 907
Db 107351 CATTGAGGAATGGTGAGAAATGAGCAATTTCTGGGTCAATTTGGTGTATGATCAGCACATCT 107410
QY 907 uPheAlaValValGlnGlyLeuLysValLeuAlaGlyIleAspThrAsnPheThrVa 927
Db 107411 CTTTCTGTGCATACAAAGGCTTCTGAAAGTTCTTCTGTAATTTGACAAACTTCACTGT 107470
QY 927 tThrSerLysAlaThrGlyAspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTr 947
Db 107471 TACATCCAAAGGCAACA-----GATGATGAGGAATTTGAGAAATTTGACCCATCAAGTG 107524
QY 947 pThrThrLeuLeuIleProProThrThrLeuLeuLeuIleLeuValIleGlyValValAl 967
Db 107525 GACCACTCTCTTAATTTCTCCAACGACTATCTTAATAATCAATATTTGTTGGGTTGTTGC 107584
QY 967 aGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTyrGlyProLeuPheGlyLysLe 987
Db 107585 TGGAAATCTCAGAGCCCAATTAACATGTTTCCATCATGGGACCTCTATTTTGGAAAACT 107644
QY 987 uPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyAr 1007
Db 107645 CTTCTTTTCTTCTGGGTGATTTGCCATCTATATCCATTTCTTAAAGGTTTGTATGGTCTG 107704
QY 1007 gGlnAsnArgThrProThrValValValIleTrpSerIleLeuLeuAlaSerIlePheSe 1027
Db 107705 GCAAAACCGCACCCCCACCATTTGTTGATATGTTGTCAGTGTGTTGGCTTCCATTTTCTC 107764
QY 1027 rLeuLeuTrpValArgIleAspProPheIleValArgThrLysGlyProAspValArgGl 1047
Db 107765 TTTGTTATGGGTAGAAATTTGACCCATTCGTATGATAAACTAAGGACCTGATACCAAGCT 107824
QY 1047 nCysGlyIleAsnCys 1052
Db 107825 ATGCGAATCAACTGT 107840

RESULT 14
AX048245
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .7234
/organism="Arabidopsis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:29726"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 7234
Score: 4224.50 Matches: 856
Percent Similarity: 61.35% Conservative: 87
Best Local Similarity: 55.69% Mismatches: 90
Query Match: 74.80% Indels: 508
DB: 6 Gaps: 20

US-10-627-132-30 (1-1052) x AX048245 (1-7234)
QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
Db 92 ATGGAAGCTAGCGCGGCTTTGTGCGCGGTTCTCATAAACCGTAATGAATAGTCTGCTATT 151
```

QY 21 ArgGlyHisGluAsp----- 25
Db 152 CACAACCATGAAGAGGTTTTTTCACATTTACTTTTTTCTCATCACTTACCAGAAAGTTTTTTTT 211
QY 25 ----- 25
Db 212 TACCAACGCTAGTAAATAATTAATTGTCATTTTTTCGTTTTATTGTTACTATTCTTAA 271
QY 25 ----- 25
Db 272 AATTCGTATGTTTGGGAATAAAATATATGATCAATTTTTTTAAATCAATCTTCTTATTATGA 331
QY 25 ----- 25
Db 332 GACAAAAATTTATAATCTGTATCTGTATCTGTAGTTGCAATAATGTTGTAGAAAAATTCATATCTT 391
QY 25 ----- 25
Db 392 TGTAGCAACATATAATTAATTTTGTGTTGTTAATTAATTAAGTTGAGAAGTCAGGTTTAAACCAT 451
QY 25 ----- 25
Db 452 TTAATCGCTGTCATTTTTTTTATTATCTTTACTTCAATAGAACTCTTTTTTTGTTGATTAG 511
QY 26 -----Pro 26
Db 512 AAATTTACAGGTTTTTATTTTCGCTTTTAAAGAACTTAAACTTTTTTGTGTAATTAACAGCCA 571
QY 27 LysProLeuArgAlaLeuSerGlyGlnValCysGluIleCysGlyAspGluValGlyLeu 46
Db 572 AAGCCTCTGAAGAATCTAGATGGACAAATCTGTGAGATATGTGAGATCAGATGGTTA 631
QY 47 ThrValAspGlyAspLeuPheValAlaCysAsnGluCysGlyPheProValCysArgPro 66
Db 632 ACAGTAGAAGGAGACCTCTTCGTAGCTTGCATGAGTGTGTTTTTCCGGCGTAGACCT 691
QY 67 CysTyrGluTyrGluArgGluGlyThrGlnAsnCysProGlnCysLysPheThrArgTyr 86
Db 692 TGCTATGATGACGAGAAGAAGAAACACAAACTCTCTCAGTGTAAAGACTCGTTAC 751
QY 87 LysArgLeuLys----- 90
Db 752 AAGGCTCTCAGAGGTAAAGTTATTATTATCTCCCTCTGCTCTTGTGTTGTCGACGAA 811
QY 90 ----- 90
Db 812 TGCCTCTATGAAATTTAAAAAGGCTGTTCTTTTTTTTGTGAACTTGAGAGTAATG 871
QY 91 -----GlySerProArgValAlaGlyAspAspGluAsp 103
Db 872 ATCTGTTTTTTTGGTTTCTGAAGGAAGCCCAAGAGTGGAGGAGATGAAGACGAAGAAGAT 931
QY 104 IleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnArgGlnLeuGlu 123
Db 932 ATTGATGATATTGATGATGAATTAATTAATTCGAACATGAACAGAT----- 976
QY 124 GlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMetSerTyrGly 143
Db 977 -----AAGCATAAGCATCTCTGCTGAGGCTATGCTTTATGAGAAATGAGCTATGGA 1027
QY 144 ArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIleProProIleIleThrGly 163
Db 1028 AGAGGTCTCGAGGATCATGAATATGCG-----AGATTCCCACTGTTATAGCTGCT 1078
QY 164 SerArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrGlyHisGlyGlu 183
Db 1079 GGTCA-----AGTGGAGATTTCCAGTTGGAGGAGTTATGTTAATGGAGAA 1126
QY 184 ValSerSerLeuHisLysArgIleHisProTyrProValSerGluProGlySerAla 203
Db 1127 -----CATGGGCTTCATAAGCGTGTGCACCCATATCCATCATCTCAAGCTGGTGAAGTCT 1180

QY 204 LysTrpAsp-----GluLysLysGluValSer----- 212
Db 1181 CA-TGGAAATGTTAACTTACATATAGATTTTAAGAAATGCTCTCACAGTGATGATTAGTAGG 1239
QY 212 ----- 212
Db 1240 GTCATGCATATCTCCATATGTGCAATAAACAATAGTATGAGGCTTCCAGCTTAAATAGTA 1299
QY 212 ----- 212
Db 1300 GATAGGACATAGTTTTCATAAACATGGACTTTGGGTTCTATTACATCTTTCTTATGAAAT 1359
QY 212 ----- 212
Db 1360 TCATCAGCAGACCTTTTCTTAACAATACTTCTCTTTTGTGTTATGTTGCTAAATTTTAAAT 1419
QY 213 -----TrpLysGluArgMetAspAspTrpLysSerLysGlnGly 225
Db 1420 GTGTTAGGAGTGTAGGAGGATGCGGGAAGAATGAGTACTGGAAGCTCCAGCATGGA 1479
QY 226 IleLeuGlyGlyGlyAlaAspProGlu---AspMet-AspAlaAspValAlaLeuAsn-- 243
Db 1480 AATCTTGGGCGAGAACCCAGATGATGATCCTGAGATGGAGCTGTAATGCCCTCCACAAACAT 1539
QY 243 ----- 243
Db 1540 TTATCTAAGACATCAGTTTTTGTATGATTGTTGATCTTACAAAATTTTGGATTGAC 1599
QY 244 -----AspGluAlaArgGlnProLeuSerArgLysValSerIleAl 257
Db 1600 TGGAAATGGCTGTAGGATCGAGGACGCGCACCTCTCGCGGAAGTTCCTCATGTC 1659
QY 257 aSerSerLysValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPh 277
Db 1660 CTCGAAGCAATCAATCATATCGATGGTGCATGTTGCTAGGCTTGTGATCTTAGCAGT 1719
QY 277 ePheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValse 297
Db 1720 TTTTTCGGTATAGGCTCTTGAATCCAGTGCATGATGCTCTGGGATTATGGCTGACCTC 1779
QY 297 rIleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPh 317
Db 1780 TGTGATCTGTGAATCTGTTGCTGCTCTCTGATTCCTGATCAGTTCCTCCCAAGTGGTT 1839
QY 317 eProIleAspArgGluThrTyrLeuAspArgLeuSerLeu----- 330
Db 1840 CCTTATTGAACGTGAGACCTATCTAGATCGGCTTTCCTCAGGTAAATCCACAGATTCT 1899
QY 330 ----- 330
Db 1900 CAAGTAGAAGTCTTAAATCTATGACGTTGGAGTTTGGATGTAATAATTTTTTGTATTATA 1959
QY 331 -ArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSerTh 350
Db 1960 CAGGTACGAGAGAGAGTGAACCAATATGCTTGGCCCTGTAGATGCTCTTGTGCACTAC 2019
QY 350 rValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAl 370
Db 2020 GGTGACCCCATGGAAGAGGCTCCCTCGTCGTACATCAACACTGTGCTGTCAATCTGGC 2079
QY 370 aValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLe 390
Db 2080 CATGACTTACCAGTTGAGAAATCTCTGTATGCTCTCTGACGCGTGTCTCAATGCT 2139
QY 390 uThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLys 410
Db 2140 TACATTCGAATCTCTCTCGAAACTGCTGAGTTTGCAGAAAAATGGGTTCTCTCTGTA 2199
QY 410 sLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLe 430
Db 2200 GAAATTTCTCCATAGACCCACGGGACCGGAGATGTACTTTCAGTTGAAGTTGATTACT 2259
QY 430 uLysAspLysValGlnProThrPheValGlnGluArgAlaMetLys----- 446

Qy 997 euYrProPheLeuLeuGlyLeuMetGlyArgGlnAsnArgThrProThrValValVali 1017
 Db 4471 TCTACCACTTCTCTCAAGGCTGATGGGTAGACAGAACAGACACCAACCACTTGTGTGA 4530
 Qy 1017 leTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTTPValArgIleAspProPheI 1037
 Db 4531 TTTGGTCAGTGTATATGGCATCTATCTCTCTTCTTGGTAAAGAAATGATCCCTTTG 4590
 Qy 1037 leValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
 Db 4591 TGCTCAAGACCAAGGACCTGACACTTCCAAGTGTGCATCAACTGC 4637

RESULT 15
 AF091713 7234 bp DNA linear PLN 27-JUN-1999
 LOCUS Arabidopsis thaliana cellulose synthase catalytic subunit (IRX3)
 DEFINITION gene, complete cds.

ACCESSION AF091713 GI:5230422
 VERSION AF091713.1
 KEYWORDS Arabidopsis thaliana (thale cress)
 SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 1 (bases 1 to 7234)
 Taylor,N.G., Scheible,W.R., Cutler,S., Somerville,C.R. and
 Turner,S.R.

TITLE The irregular xylem3 locus of Arabidopsis encodes a cellulose
 synthase required for secondary cell wall synthesis

JOURNAL Plant Cell 11 (5), 769-780 (1999)
 MEDLINE 99264300
 PUBMED 10330464
 REFERENCE 2 (bases 1 to 7234)
 AUTHORS Taylor,N.G. and Turner,S.R.
 TITLE Direct Submission
 JOURNAL Submitted (09-SEP-1998) Biological Sciences, 3,614 Stopford
 Building, University Of Manchester, Oxford Road, Manchester M13
 9PT, UK

FEATURES

source

1..7234
 /organism="Arabidopsis thaliana"
 /mol_type="Genomic DNA"
 /cultivar="Landsberg erecta"
 /db_xref="taxon:3702"
 /chromosome="v"
 <92..>4640
 /gene="IRX3"
 /note="irregular xylem3"
 join(<92..166,569..764,894..1172,1427..1520,1615..1881,
 1963..2308,2396..2659,2737..2949,3029..3242,3338..3534,
 3625..3978,4059..>4640)
 /gene="IRX3"
 /product="cellulose synthase catalytic subunit"
 join(92..166,569..764,894..1172,1427..1520,1615..1881,
 1963..2308,2396..2659,2737..2949,3029..3242,3338..3534,
 3625..3978,4059..4640)
 /gene="IRX3"
 /codon_start=1
 /product="cellulose synthase catalytic subunit"
 /protein_id="AAD40885.1"
 /db_xref="GI:5230423"
 /translation="MEASGLVAGSHNRELIVIHNEEPKPLKNLDQGFCEICDQI
 GLTVGDLFVACNECGFPACRCPCEYERREGTQNCPCQKTRYKRLRGSPRVEGDEE
 DIDDLEYFENIEHEDKHKSAEAMLYGKMSYGRGPEDDENGRRFPVITAGHSGEFV
 GGGVNGEHLKRVHPVPSSAGSEGGWRERMDWKLOHGNLGPEDDPEMGLIDE
 AKQPLSRKVPKASSKINRYMVIVARLVILAVFLRYLLNPVHDALGLWLTSVCEIW
 FAVSLVDQFPFIRETYLDRLSLRYEREPEPNMLAPVDVFSVTVDPLKEPPLVT
 SNTVLILAMDYVPSKISCYSDSDGASMLTFESSETAFARKWVPFCCKFSIEMRAP
 ENYFTLVLDYQKVPFTFKERRAMKEYEFEKVRINAQAKASKVPLEKGIWQDGT
 PWPGNNTKDHQPMIOVFLHSGGFVGEHGLPRLVYVSREKRPQFQHKKAGAMNALV
 RVAGVLITWAPFNLDCDHYVNSKAVREAMCFILMDPQIGKKVCVVPFQRPDGDITN

gene

mRNA

CDS

ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 7234
 Score: 4224.50 Matches: 856
 Percent Similarity: 61.35% Conservative: 87
 Best Local Similarity: 55.69% Mismatches: 90
 Query Match: 74.80% Indels: 508
 DB: 8 Gaps: 20
 US-10-627-132-30 (1-1052) x AF091713 (1-7234)
 Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
 Db 92 ATGGAAGCTAGCCCGCTCTTTCGCCGGTCTCTATAACCGTAATGAACCTAGTCGCA 151
 Qy 21 ArgGlyHisGluAsp----- 25
 Db 152 CACAAACCACTAGAGAGGTTTTCACATTACTTTTCTCATCACTTACCAAGTTTTTTT 211
 Qy 25 ----- 25
 Db 212 TACCACGCTAGTAAATAATATTGCACTTTTTCGTTTTATTGGTTACTATTCTTAAA 271
 Qy 25 ----- 25
 Db 272 AATTCGTATGTTTGGGAATAAAATATGATCATATTTTAAATCATCTCTTTATATGA 331
 Qy 25 ----- 25
 Db 332 GACAAAATTTATAATCTGTATTCTGTAGTTGCAATAATGTTGAGAAAATTCATATCT 391
 Qy 25 ----- 25
 Db 392 TGTTAGCAACATAATAATTTTGTGTAATAATTAAGTTGAGAAGTCAGGTTTAACCAT 451
 Qy 25 ----- 25
 Db 452 TTAATCGCTGTCACTTTTTTTTATATCTTTACTTATAGAACTTTTTTTTGTGATTAG 511
 Qy 26 -----Pro 26
 Db 512 AATTTTCAGGTTTTATTTCGTTTAAAGAACTTTAAACTTTTGTGTAATAATAACAGCCA 571
 Qy 27 LysProLeuArgAlaLeuSerGlyGlnValCysGluLeuCysGlyAspGluValGlyLeu 46
 Db 572 AAGCCTCTGAAGAATCTAGATGGCAATTTCTGTGAGATATGTGGAGATCAATTTGGTTA 631
 Qy 47 ThrValAspGlyAspLeuPheValAlaCysAsnGluCysGlyPheProValCysArgPro 66
 Db 632 ACAGTAGAAGGAGACCTCTTCGTAGCTTGCATGAGTGTGTTTCCGGCGGTAGACCT 691
 Qy 67 CysTyrGluTyrGluArgArgGluGlyThrGlnAsnCysProGlnCysLysThrArgTyr 86
 Db 692 TGCTATGATGACGAG 751
 Qy 87 LysArgLeuLys----- 90
 Db 752 AAGCGTCTCAGAGGTAAGTTATTTATTAATCTCCCTCTGCTCTTGTGTGTTGTCGACG 811
 Qy 90 ----- 90
 Db 812 TGCCTCTATGAAATTTAAAGGCTGTCTCTTTTATTAGTTGAACTTGAGAGTAATG 871
 Qy 91 -----GlySerProArgValAlaGlyAspAspGluGluAsp 103

872 ATCTGTTTTCGTTCTCAAGAGCCCAAGAGTGGAGGAGATGAAGACCAAGAGAT 931
104 IleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnArgGlnLeuGlu 123
932 ATTGATGATATTGAGTATCAATTAATATCGAACATGAACAAGAT----- 976
124 GlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMetSerTyrGly 143
977 -----AAGCATAAAGCATTTCTGCTGAGGCTATGCTTTATGGGAAATGAGCTATGGA 1027
144 ArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIleProProIleIleThrGly 163
1028 AAGAGTCTGAGATGATGAGATGG-----AGATTCCACCTGTTATAGCTGTT 1078
164 SerArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrGlyHisGlyGlu 183
1079 GGTGAT-----AGTGGAGAAATTTCCAGTTGGAGGAGTTATGTTAATGGAGAA 1126
184 ValSerSerLeuHisGlyArgIleHisProTyrProValSerGluProGlySerAla 203
1127 -----CATGGCTTTTCAAGCGTGTGCACCCATATCCATCATCTGAAGCTGGTGTCT 1180
204 LysTyrAsp-----GluLysLysGluValSer----- 212
1181 CA-TGGAATGTTAACTTACATATAGATTTAAGAATGCTCTCACAGTATGATTAGT 1239
212 ----- 212
1240 GTCATGCATATCTCCATATGTGCAATAACATAAGTATGAGCGCTTCAGCTTAATAGTA 1299
212 ----- 212
1300 GATAGGGACATAGTTTCAATAACATGAGCTTTGGGTTCTATTACATCTTCTATGAAAT 1359
212 ----- 212
1360 TCATCAGCAGACCTTTTCTAACAAATCTTTCTTTTGTGTTATGTTGTTAATTTAAT 1419
213 -----TrrLysGluArgMetAspAspTrrLysSerLysGlnGly 225
1420 GTGGTAGGAGTGGAGGAGTGGCGGAAAGAAATGGATGACTGGAGCTCCAGCATGGA 1479
226 IleLeuGlyGlyAlaAspProGlu---AspMet-AspAlaAspValAlaLeuAsn--- 243
1480 AATCTTGGCCAGAACCAAGATGATGATCTCGAGATGGGACTGTAATGCTCCACAAACAT 1539
243 ----- 243
1540 TTATCTAAGACATCAGTTTGTATGATTGGATTTCATGCTTACAAAAATTTTGGATTGAC 1599
244 -----AspGluAlaArgGlnProLeuSerArgLysValSerIleAl 257
1600 TGGAAATGGCTGTAGATCGACGAGCAGCGGACCACTCTCGCGGAAAGTTCCCATTTGC 1659
257 sSerSerLysValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPh 277
1660 CTCGAAGCAAGATCAATCCATATCGGATGCTCATCTGCTAGGCTTGTGATTCTAGCAGT 1719
277 ePheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrrProValIse 297
1720 TTTTCTGGGATAGGCTCTTGAATCCAGTGCATGCTCTGCGGATTTATGGCTGACCTC 1779
297 rIleIleCysGluIleTrrPheAlaIleSerTrrIleLeuAspGlnPheProLysTrrPh 317
1780 TGTGATCTGTGAATCTGGTTCGCTGCTCTCTTGGATTCTTGATCAGTTCCCCAGTGGTT 1839
317 eProIleAspArgGluThrTyrLeuAspArgLeuSerLeu----- 330
1840 CCTATTGAACGTGAGACCTATCTAGATCGGCTTTTCCCTCAGGTAAAAATCCACAGATTCT 1899
330 ----- 330

1900 CAACTAGAAAGTCTTAAAAATCTATGACGTTGGAGTTTGGATGAAATATATTTTGTATTATA 1959
331 -ArgTyrGluArgGluGlyGluProSerSerLeuLeuSerAlaValAspLeuPheValSerTh 350
1960 CAGGTACGAGAGAGAGAGTGAACCAATATGCTTCCCTCTGTAGATGCTTTTGTCAATAC 2019
350 rValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAl 370
2020 GGTGACCATTCATGAAGGAGCTCCCTCTGTCATCATCAACATGCTGCTCAATCTTGC 2079
370 aValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetIle 390
2080 CATGATACCATCCAGTTGAGAAATCTCTGCTATGCTCTGACGAGGCTTCAATGCT 2139
390 uThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrrPheValProPheCysLy 410
2140 TACATTCGAATCTCTCTCGAAACTGCTGAGTTTGCAGAGAAATCGGTTCCCTTCTGTAA 2199
410 sLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrIle 430
2200 GAAATTTCTCCATAGAGCCACGCGGACCTGACTTCTGCTGTTGAAAGTTGATTATCT 2259
430 uLysAspLysValGlnProThrPheValGlnGluArgAlaMetLys----- 446
2260 CAGGACAAAGTCCACCAACATTTGTTAAGGAAGCTGAGCCATGAAGGTCAGTGTATA 2319
446 ----- 446
2320 TCACCTGATCTAGTTATATACCAACCCCATCTTTTCACTATAATCTAACTTCATAGTGA 2379
447 -----ArgGluTyrGluGluPheLysValArgIleAsnAlaLeuValAl 461
2380 AATTGACATATCAGAGAGAGATATGAGGAGTTCAAGGTCAGGATCAACGCTCAAGTGC 2439
461 aLysAlaMetLysValProAlaGluGlyTrrIleMetLysAspGlyThrProTrrProGl 481
2440 GAAGCCCTCAAGGTTCTCTAGAGGTTGATCATGCAAGATGAAACACCGTGGCCAGG 2499
481 YAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHi 501
2500 GAACAAACCAAGGACCAACCCCGGTATGATCCAAAGTCTTCTCGGCCACACGCGAGGATT 2559
501 sAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGl 521
2560 TGATGTCAAGGCGATGAGCTTCTCGGCTTGTGTACGTGTCCGTGAGAAGCGTCTGG 2619
521 yPheGlnHisIleLysLysValGlyAlaMetAsnAlaLeuIle----- 535
2620 TTTTCAACACCAAGAAAGCTGGCGCATGAATGCCCTGGT-AATTTTCTTGTATCTGCC 2678
535 ----- 535
2679 TTGGACCAACCAAGAAATGATCTCCGTGTCTTGGACCTTAACCTGATACTTCTGTCAAGT 2738
536 -ArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTy 555
2739 TCGATGTGCGAGCGTACTCACAATGCTCTTTCATGCTGAACCTGGACTGTGATCACTA 2798
555 rIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGl 575
2799 TGTAAACCAACAGCAAGCGCGTGGAGGAAGCAATGTGTTTGTGATGGATCTCCAGATTGG 2858
575 YArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspAr 595
2859 AAAGAGGTCTGCTATGTTTCAAGTTCCCTCAAGGTTTGTATGGCATTTGACACAAACGATCG 2918
595 gTyrAlaAsnArgAsnThrValPhePheAsp----- 605
2919 TTACGCCCAACAGAAACACAGTCTTCTTTTGTGTAAGACTCAATTCTATATTTTCCAACCT 2978
606 -----IleAsnMetL 609
2979 CTGGTCTAAATGAAATGTATACCTCTTGTCTTACACTCTTGTGTTGCTACAGATCAATATGA 3038


```
QY 609 ySgLyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgG 629
Db 3039 AAGGCTAGATGGAAATCCAAAGTCCAGTTTACGTTGGTACTGTTGGTTTTCACAGCAC 3098
QY 629 lnAlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysA 649
Db 3099 AAGCTCTGTATGTTATGAACCAACCAAGGGTCTCAAAAGCTCCAAAGATGATAAGCTGTG 3158
QY 649 spCysCysProCysPheGly-----ArgLysLysArgLysHisAlaLysAspGlyLeu- 666
Db 3159 GTTGTCTCTCTGTTTGGGCGCGGAGAAAGAAATAAGAAATTTTCCAGAATGACATGA 3218
QY 667 -----ProG 668
Db 3219 ATGGTGAGGTAGACAGCCCTGGAGGTAATAATATCCCAACCAACCTTATAATATCATGTCAT 3278
QY 668 luGlyThrAlaAspMet-----G 674
Db 3279 TCCTT-GCAGTAGATTTTCGTTTATGTTGGAATCTTTGCGGATCTGATAGTGTGTTTGGCAG 3337
QY 674 lyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyG 694
Db 3338 GAGCAGAAGGTGATAAGGACCATTTGATGTCGAAATGAATTTTGAGAAAAACATTTTGGGC 3397
QY 694 lnSerAlaAlaPheValThrSerThrLeuMetGluGlyGlyValProProSerSerS 714
Db 3398 AATCATCATCTTTGTAACTCAACTTTGATGGAAGAGGGTGGTTCCTCCGTCATCAA 3457
QY 714 erProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysT 734
Db 3458 GTCTTCAGTGTCTCTTAAAGAGGCAATCCATGTCATAAGCTGCGGTTATGAAGACAAGA 3517
QY 734 hrAspTyrGly----- 737
Db 3518 CTGAGTGGGAACTCAGGTAATAATACTGAATCGTAGAAATCACCTTCTTATTGTGATT 3577
QY 738 -----LeuGluLeuGlyTyrIleT 744
Db 3578 TAGTAGTGGCCATTGTAAAAAAGCTTTGTGTATCTCGAAATTCGAGCTGGTTGGATCT 3637
QY 744 yrgLysIleThrGluAspIleLeuThrGlyPheLysMetHisCysArgGlyTyrArgS 764
Db 3638 ATGGCTCTATCACAGAGGATATTTTGACGGGATTCAGATGCATTTGCCGTGGATGGAGGT 3697
QY 764 erValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIleAsnLeuSerA 784
Db 3698 CTATTTACTGTCATGCTTAAGAGGCTGCATTCAAAGGTTTCAGCTCCTATTATTAATCTATCAG 3757
QY 784 spArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIlePhePheSerArgH 804
Db 3758 ACAGGTTAAACAGGTTTTTGGGTGGGCACCTTGGATCGGTTGAGATATTTTTCAGCCGGC 3817
QY 804 isSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTyrLeuGluArgPheAlaT 824
Db 3818 ACAGTCTCTCTGTTATGGCTACAAAGGAGGCAAACTCAAGTGGCTTCAGCGTTTTGCTT 3877
QY 824 yrlleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyrCysThrLeuP 844
Db 3878 ATGCCAACACACAACATCTACCCCTTCACATCTATACCACTTCTTGGCTTACTGTATCCTTC 3937
QY 844 roAlaValCysLeuLeuThrGlyLysPheIleMetProSer----- 857
Db 3938 CAGCCATCTGCTCTCTTACTGACAAATTCATCATGCCACC-GGTTAGTAAAAATATATCAGA 3996
QY 857 ----- 857
Db 3997 GAAAGACACTTAGAAGCTGCATCAAAATGTGCTAACTATCTGTTTTTCCCAATTTTCTTTTC 4056
QY 858 --IleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrG 877
Db 4057 AGATAAGCACATTTTGTAGTCTCTTCTTCTCATCTCTGTTTATGTCGATCATTTGTAACGG 4116
```

```
QY 877 lyIleLeuGluMetArgTyrTrpSerGlyValSerIleGluTyrTrpArgAsnGluGlnP 897
Db 4117 GAATCTTTGAAATTGAGATGGAGCGGAGTTAGCATTTGAGGAATGCTGGAGAAACGAGCAAT 4176
QY 897 heTyrValIleGlyGlyValSerAlaHisIleuPheAlaValValGlnGlyLeuLeuYev 917
Db 4177 TCTGGGTCTATTGGAGGAATCTCAGCTCATCTCTTTCGCGTGTGTCGAAGGTCTCCTCAAAA 4236
QY 917 alLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspA 937
Db 4237 TCTTACGAGCATTTGACAAACTTCCCGTTCATCAAAAGGCAACA-----GATGATG 4290
QY 937 spGluPheAlaGluLeuTyrAlaPheLysTyrThrThrLeuLeuIleProProThrThrL 957
Db 4291 ATGACTTTTGAGAACTTTTACGCATTCAAATGGCAACACTGCTGATCCCTCCCAACAAC 4350
QY 957 euLeuIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyT 977
Db 4351 TCTTAATCATAAACATTTGTGGCGTTGTTGCGAGGCATCTCAGATGCCATTAACAATGGAT 4410
QY 977 yrgLysSerTyrGlyProLeuPheGlyLysLeuPhePheAlaPheTyrValIleValHisL 997
Db 4411 ATCAGTCTTGGGAGCCTCTATTGTTGGTAACCTCTTCTTCTCTTGGTTCATTTGCATC 4470
QY 997 euTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValI 1017
Db 4471 TCTACCCATTTCTCTCAAGGTCCTGATGGGTAGACAGAACACCAACCAATTTGTGTGA 4530
QY 1017 leTyrSerIleLeuLeuAlaSerIlePheSerLeuLeuTyrValArgIleAspProPheI 1037
Db 4531 TTTGGTTCAGTGTATTGCGCATCTATCTCTTCTTGTGTTGGGTAAAGAAATGTATCTTTTG 4590
QY 1037 leValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 4591 TGCTCAAGACCAAGGACCTGCACCTTCCAAGTGTGGCATCAACTGC 4637
```

Search completed: March 12, 2005, 13:08:15

Job time : 7860 secs